

STIC-Biotech/ChemLib

188683

From: Whiteman, Brian  
Sent: Thursday, May 04, 2006 12:56 PM  
To: STIC-Biotech/ChemLib  
Subject: FW: seq search

10659800

SEQ ID NO 6 against AA and DNA

- 1) issued us patents and published us patent applications
- 2) **commercial databases**

**FOR DNA limit to 500 nucleotides or less.**

Thank you,

Brian Whiteman  
Remsen, 2D14  
mail box 2C18  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
(571) 272-0764

CRFE

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

- Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).
- Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**THIS PAGE BLANK (USPTO)**



## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2006, 12:28:10 ; Search time 231 Seconds  
(without alignments)  
1490.467 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MGDRGSSRRRTGSRPSSHG.....VLMYVHDYVLYNTEAPAAEA 488

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2588	99.8	488	1 DGAT1_HUMAN	O75907 homo sapien
2	2504.5	96.5	491	1 DGAT1_CERAE	O9gmfi cercopithe
3	2299.5	88.6	500	2 OEBH15_RAT	O8bh15 rattus norv
4	2292.5	88.4	489	2 O8SQB0_BOVIN	O8ebd0 bos taurus
5	2279.5	87.9	489	2 O8MK44_BOVIN	O8mk44 bos taurus
6	2278.5	87.8	498	1 DGAT1_MOUSE	O922a7 mus muscul
7	2278.5	87.8	498	2 O54A46_MOUSE	O54a46 mus muscul
8	2266.5	87.4	489	2 O8MHZ1_PIG	O8mh21 sus scrofa
9	2266.5	87.4	498	1 DGAT1_RAT	O9erm3 rattus norv
10	2117	81.6	400	2 O9BRH5_HUMAN	O9bdh5 homo sapien
11	1686	65.0	499	2 O6P3J0_BRARE	O6p3j0 brachydanid
12	1648.5	63.6	507	2 O6DHG2_BRARE	O6dhg2 brachydanid
13	1624	62.0	473	2 O4RLB7_TESTNG	O4rlb7 tetraodon r
14	1608	62.0	447	2 O4RL18_TESTNG	O4rl18 tetraodon r
15	1230	47.4	242	2 O91YB5_RAT	O91yb5 rattus norv
16	1122.5	43.3	497	2 O60MK6_CABER	O60mk6 caenorhabd
17	1106	42.6	496	2 O45245_CABEL	O45245 caenorhabd
18	1106	42.6	498	2 O9NCE1_CABEL	O9nce1 caenorhabd
15-4-1102.5	42.5	480	2 O7Q3B6_ANOGA	O7q3b6 anopheles g	
20	1018	39.2	565	2 O960U8_DROME	O960u8 drosophila
21	1018	39.2	615	2 O7KTS5_DROME	O7kts5 drosophila
22	1016	39.2	565	2 O8ST50_DROME	O8st50 drosophila
23	975	37.6	221	2 O4VYU5_BUBBU	O4vyu5 budaius bu
24	823.5	31.7	498	2 O5GKZ7_SOYBN	O5gkz7 glycine max
25	821	31.6	655	2 O55BH9_DICDI	O55bh9 dictyosteli
26	819.5	31.6	511	2 O512P3_LOTUA	O512p3 lotus japon
27	819	31.6	507	2 O5UEW2_BRAUT	O5uem2 bruceus al
28	817	31.5	503	2 O4U3T8_BRAUT	O4u3t8 brassica ju
29	800	30.8	406	2 O6DNG7_SOYBN	O6dng7 glycine max
30	797	30.7	532	2 O9SEB9_TOBAC	O9seb9 nicotiana t
31	791.5	30.5	520	2 O9SLD2_ARATH	O9slid2 arabidopsi

32	788.5	30.4	501	2 O9W4V2_BRANA	O9w4v2 brassica na
33	785.5	30.3	518	2 O8RX96_TROMA	O8rx96 tropaneolun
34	785.5	30.3	534	2 O9FUI6_PEFIF	O9fui6 perilla fru
35	784	30.2	503	2 O4U3T9_BRAUT	O4u3t9 brassica ju
36	781.5	30.1	521	2 O67C39_RICCO	O67c39 ricinus com
37	779.5	30.1	503	2 O9XGR5_BRANA	O9xgr5 brassica na
38	774.5	29.9	477	2 O52822_ORYSA	O52822 oryza sativ
39	767	29.6	538	2 O51396_ORYSA	O51396 oryza sativ
40	759	29.3	532	2 O6ED63_ORLEU	O6ed63 olea europ
41	656.5	25.3	341	2 O9XGV4_BRANA	O9xgv4 brassica na
42	639	24.6	504	2 O688L6_ORYSA	O688l6 oryza sativ
43	585	22.6	508	2 O4X1Y1_ASPEFU	O4x1y1 aspergillus
44	584	22.5	526	2 O6C9V5_YARLI	O6c9v5 yarrowia li
45	583	22.5	540	2 O7YXK9_TOXGO	O7yxk9 toxoplasma

## ALIGNMENTS

## RESULT 1

ID DGAT1\_HUMAN STANDARD; PRT; 488 AA.  
AC O75907; O96BB8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Diacylglycerol O-acyltransferase 1 (EC 2.3.1.20) (Diacylglyceride  
acyltransferase) (ACAT-related gene product 1).  
GN Name=DGAT1; Synonyms=AGRP1, DGAT1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=98434592; PubMed=9756920; DOI=10.1074/jbc.273.41.26765;  
RT Oelkers P., Behari A., Cromley D., Billheimer J.T., Sturley S.L.;  
RT "Characterization of two human genes encoding acyl coenzyme  
RT A:cholesterol acyltransferase-related enzymes.";  
RL J. Biol. Chem. 273:26765-26771(1998).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Yamasaki Y., Watanabe T.K., Tanigami A.;  
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Brain, and Skin;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RT Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang X., Hsieh F.,  
RA Diachenko L., Martina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinini P., Prange C.,  
RA Rahn S.S., Loquellano N.A., Peters G.J., Abrahamson R.D., Mullaly S.J.,  
RA Boesek S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.W., Gay L.J., Hulys S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fajlan J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Catalyzes the terminal and only committed step in  
CC triacylglycerol synthesis by using diacylglycerol and fatty acyl  
CC CoA as substrates.  
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + 1,2-diacylglycerol = CoA +

```

CC triacylglycerol.
CC -1- PATHWAY: Central role in the metabolism of cellular diacylglycerol
CC lipid.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (By similarity).
CC -1- SIMILARITY: Belongs to the membrane-bound acyltransferase family.
CC Sterol o-acyltransferase subfamily.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL, AF059202; AAC63997.1; -; mRNA.
DR EMBL, AB057815; BAC66170.1; -; mRNA.
DR EMBL, BC015762; AAH15762.1; -; mRNA.
DR EMBL, BC023565; AAH23565.1; -; mRNA.
DR HGNC, HGNC:2843; DGAT1.
DR H-InvDB, HIX0007854; -.
DR Reactome, O75907; -.
DR MIM, 604900; -.
DR GO, GO:0008415; F:acyltransferase activity; TAS.
DR GO, GO:0004144; F:diacylglycerol O-acyltransferase activity; TAS.
DR GO, GO:0006641; P:triacylglycerol metabolism; TAS.
DR InterPro, IPR004299; MBOAT_fam.
DR Pfam, PF03062; MBOAT, 1.
KW Acyltransferase; Endoplasmic reticulum; Transferase; Transmembrane.
FT TRANSMEM 104 124
FT TRANSMEM 130 150
FT TRANSMEM 166 186
FT TRANSMEM 189 209
FT TRANSMEM 282 302
FT TRANSMEM 332 352
FT TRANSMEM 406 426
FT TRANSMEM 428 448
FT TRANSMEM 453 473
FT ACT SITE 415 415
FT CONFLICT 129 129 Y -> H (in Ref. 1).
SQ SEQUENCE 488 AA; 55278 MW; 6574D5DBF15D6171 CRC64;
Query Match 99.8%; Score 2588; DB 1; Length 488;
Best Local Similarity 99.8%; Pred. No. 2.2e-199;
Matches 487; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 421 VPLRMFLMFTGMMMAQIPLAMFYGRFPGNYGNAVWLSLIIIGPILAVLMYVHDYVLN 480
DB 421 VPLRMFLMFTGMMMAQIPLAMFYGRFPGNYGNAVWLSLIIIGPILAVLMYVHDYVLN 480
OY 481 YEAPAAEA 488
DB 481 YEAPAAEA 488
-----
RESULT 2
DQAT1_CERAE STANDARD; PRT; 491 AA.
AC O9GMFL1
ID 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Diacylglycerol O-acyltransferase 1 (EC 2.3.1.20) (Diglyceride
DE acyltransferase).
GN Name=DGAT1; Synonyms=DGAT;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Cercopithecoidea; Cercopithecinae; Cercopithecus.
OC NCBI_TaxID=9534;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Liver;
RA Joyce C.W., Davis M.A., Anderson R.A., Rudel L.L.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Catalyzes the terminal and only committed step in
CC triacylglycerol synthesis by using diacylglycerol and fatty acyl
CC CoA as substrates.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1,2-diacylglycerol = CoA +
CC triacylglycerol.
CC -1- PATHWAY: Central role in the metabolism of cellular diacylglycerol
CC lipid.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (By similarity).
CC -1- SIMILARITY: Belongs to the membrane-bound acyltransferase family.
CC Sterol o-acyltransferase subfamily.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL, AF236018; AAF98557.1; -; mRNA.
DR InterPro, IPR004299; MBOAT_fam.
KW Acyltransferase; Endoplasmic reticulum; Transferase; Transmembrane.
FT TRANSMEM 107 127
FT TRANSMEM 131 151
FT TRANSMEM 169 189
FT TRANSMEM 192 212
FT TRANSMEM 285 305
FT TRANSMEM 335 355
FT TRANSMEM 404 424
FT TRANSMEM 431 451
FT TRANSMEM 456 476
FT ACT SITE 418 418
SQ SEQUENCE 491 AA; 55644 MW; BFD3683453D588DB CRC64;
Query Match 96.5%; Score 2504.5; DB 1; Length 491;
Best Local Similarity 95.9%; Pred. No. 1.2e-197;
Matches 471; Conservative 8; Mismatches 9; Indels 3; Gaps 1;

```

```

Query Match Similarity      88.6%; Score 2299.5; DB 2; Length 500;
Best Local Similarity      85.9%; Pred. No. 3,6e-176;
Matches 427; Conservative 26; Mismatches 31; Indels 13; Gaps 3

Qy      1  MGBR---GSSRRRRGSRPSSHGGCGGAPAAAEVEVDAAAGPDVGAAGDAPAPAP----- 51
Db      1  MGDREGAGASSRRRRRGSRVSIQGGSGPMDDEEVNDAAVGPDLGAGDAPAPAPVAPAAH 60
Qy      52 --MKDGDAGVSGSHNELRCHRIQDSIFSSDSGFSNRYGLINNCVWLLISNARFLENTLI 109

```

	Qy	Db	Oy
Query Match	88.4%	Score 2292.5	DB 2
Best Local Similarity	88.0%	Pred. No. 1.3e-175	
Matches	434	Conservative 15	Mismatches 31
			Indels 13
			Gaps 3
	1	MGDPR---GSSRRRTGSRFSPSHGGCGPPAAAEVRDAAAGDVGAGGAPAPAPKQDGA	57
	1	MGDGRGAGGSSRRRTGSRFSPISIQGSGGPPAAAEVYR-----DVAGGSGAPAPRDTQKGDV	54
	58	GVGSGHWEIHRCHRLQDLSLFSSDSGFSFNRYGIILNCCVMILISNARLFENLIIKTYGILVDP	117

```
Db 55 DVSGGHMDLRCHRLQDSLFSSDSGFSNVRGILNMCVWMLISNARLFLENLIKYGILVDP 114
Qy 118 IQVNSFLKDPHSPAPCLVIAANVFAVAAPQVEKRLAVGALTEOAGLLHVNATLTC 177
Db 115 IQVNSFLKDPYSWPAKCLVIAANVFAVAAPQVEKRLAVGALTEOAGLLHVNATLTC 174
Qy 178 FPAVAVLLVESITPTVGSILALMAHTILFLKLFSDYDVNSWCR-----ARAKASAGKAS 233
Db 175 FPAVAVLLVESITPTVGSILALMAHTILFLKLFSDYDVNSWCR-----ARAKASAGKAS 234
Qy 234 SAAPHTVSPYDNLTYRDLYYFLFAPTLCYELNPPSPRIKRLRLRILEMLEFTQLOV 293
Db 235 GGAQRTVSPYDNLTYRDLYYFLFAPTLCYELNPPSPRIKRLRLRILEMLEFTQLOV 294
Qy 294 GLIQOMVPTIIONSKPKMDYSRIERLKLAVPNHLIWLIFPYWLFHSCLNVAELM 353
Db 295 GLIQOMVPTIIONSKPKMDYSRIERLKLAVPNHLIWLIFPYWLFHSCLNVAELM 354
Qy 354 QGDRREFYRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMAARTAVFLASAF 413
Db 355 QGDRREFYRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMAARTAVFLASAF 414
Qy 414 FHEYLVSPLMFLMAFTGMAAQIPLAMVGRFFQGNVGNAAVWLSLIIQPIAVLMYV 473
Db 415 FHEYLVSPLMFLMAFTGMAAQIPLAMVGRFFQGNVGNAAVWLSLIIQPIAVLMYV 474
Qy 474 HDYVLYNTEAPAA 486
Db 475 HDYVLYNTEAPAA 487
```

## RESULT 5

```
08MK44_BOVIN PRT; 489 AA.
ID 08MK44_BOVIN PRELIMINARY;
AC 08MK44;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Acyl-CoA:1,2-diacylglycerol O-transferase (EC 2.3.1.20).
GN Name=Dgat;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22103627; PubMed=12077321; DOI=10.1073/pnas.142293799;
RA Witter A., Kramer W., Werner F.A.O., Kollers S., Kata S.,
RA Dusterewitz G., Bultkamp J.B., Womack J.E., Thaller G., Fries R.;
RT "Association of a lysine-232[Cys1134]alanine polymorphism in a bovine
RT gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with
RT variation at a quantitative trait locus for milk fat content.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9300-9305(2002).
DR EMBL: AJ18490; CAC86391.1; -; Genomic DNA.
DR GO: GO:0008415; F:acyltransferase activity; IEA.
DR GO: GO:0004144; F:diacylglycerol O-acyltransferase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT_1.
KW Acyltransferase; Transferase.
SQ SEQUENCE 489 AA; 55445 MW; 5C09258DDA027053 CRC64;
```

```
Query Match 87.9%; Score 2279.5; DB 2; Length 489;
Beet Local Similarity 87.6%; Pred. No. 1.4e-174;
Matches 432; Conservative 15; Mismatches 33; Indels 13; Gaps 3;
```

```
Qy 1 MGDR---SSRRRRGSRSSSHGGGPPAAEEVDAAAGPVGAAGAPAPAPKNDGA 57
Db 1 MGDRGAGGSRRRRGRSISIGSGGPPAAEEV-----DVGAGDAPVDDTKDGDV 54
Qy 58 GVGSGHWEIRCHRLQDSLFSSDSGFSNVRGILNMCVWMLISNARLFLENLIKYGILVDP 117
```

```
Db 55 DVSGGHMDLRCHRLQDSLFSSDSGFSNVRGILNMCVWMLISNARLFLENLIKYGILVDP 114
Qy 118 IQVNSFLKDPHSPAPCLVIAANVFAVAAPQVEKRLAVGALTEOAGLLHVNATLTC 177
Db 115 IQVNSFLKDPYSWPAKCLVIAANVFAVAAPQVEKRLAVGALTEOAGLLHVNATLTC 174
Qy 178 FPAVAVLLVESITPTVGSILALMAHTILFLKLFSDYDVNSWCR-----ARAKASAGKAS 233
Db 175 FPAVAVLLVESITPTVGSILALMAHTILFLKLFSDYDVNSWCR-----ARAKASAGKAS 234
Qy 234 SAAPHTVSPYDNLTYRDLYYFLFAPTLCYELNPPSPRIKRLRLRILEMLEFTQLOV 293
Db 235 GGAQRTVSPYDNLTYRDLYYFLFAPTLCYELNPPSPRIKRLRLRILEMLEFTQLOV 294
Qy 294 GLIQOMVPTIIONSKPKMDYSRIERLKLAVPNHLIWLIFPYWLFHSCLNVAELM 353
Db 295 GLIQOMVPTIIONSKPKMDYSRIERLKLAVPNHLIWLIFPYWLFHSCLNVAELM 354
Qy 354 QGDRREFYRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMAARTAVFLASAF 413
Db 355 QGDRREFYRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMAARTAVFLASAF 414
Qy 414 FHEYLVSPLMFLMAFTGMAAQIPLAMVGRFFQGNVGNAAVWLSLIIQPIAVLMYV 473
Db 415 FHEYLVSPLMFLMAFTGMAAQIPLAMVGRFFQGNVGNAAVWLSLIIQPIAVLMYV 474
Qy 474 HDYVLYNTEAPAA 486
Db 475 HDYVLYNTEAPAA 487
```

## RESULT 6

```
DGAT1_MOUSE STANDARD; PRT; 498 AA.
ID DGAT1_MOUSE STANDARD;
AC 0922A7; Q9D705;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Diacylglycerol O-acyltransferase 1 (EC 2.3.1.20) (Diacylceride
DE acyltransferase).
GN Name=Dgat1; Synonyms=Dgat;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6;
RX MEDLINE=99007259; PubMed=9789033; DOI=10.1073/pnas.95.22.13018;
RA Cases S., Smith S.J., Zheng Y.-W., Myers H.M., Lear S.R., Sande E.,
RA Novak S., Collins C., Welch C.B., Lubis A.V., Erickson S.K.,
RA Farese R.V., Jr.;
RT "Identification of a gene encoding an acyl CoA:diacylglycerol
RT acyltransferase, a key enzyme in triacylglycerol synthesis";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13018-13023(1998).
```

```
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
```

```
RC STRAIN=C57BL/6; TISSUE=Testis;
RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oseko N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Haegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schiml L.M., Kanapin A., Matsuda H., Batilov S., Beisel K.W.,
RA Dalka J.A., Bradt D., Brusic V., Chochia C., Corbani L.B., Cousins S.,
RA Balle A., Dregan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
```

RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempé C.A., Seton M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hironaka-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arkawa T., Fukuda S.,  
 RA Hara A., Hahizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 (3)  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RX MEDLINE:22388257; PubMed:12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,  
 RA Bosak S.A., Morwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Vallalon D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC - FUNCTION: Catalyzes the terminal and only committed step in  
 CC triacylglycerol synthesis by using diacylglycerol and fatty acyl  
 CC CoA as substrates.  
 CC - CATALYTIC ACTIVITY: Acyl-CoA + 1,2-diacylglycerol = CoA +  
 CC triacylglycerol.  
 CC - PATHWAY: Central role in the metabolism of cellular diacylglycerol  
 CC lipid.  
 CC - SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum.  
 CC - SIMILARITY: Belongs to the membrane-bound acyltransferase family.  
 CC Sterol o-acyltransferase subfamily.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL: AF078752; AAC72917.1; -; mRNA.  
 DR EMBL: AK008995; -; NOT\_ANNOTATED\_CDS; mRNA.  
 DR EMBL: BC003717; AA003717.1; -; mRNA.  
 DR Ensemble: ENSMUSG00000022555; Mus musculus.  
 DR MGI: MGI:133825; Dgat1.  
 DR GO: GO:0016021; C:Integral to membrane; TAS.  
 DR GO: GO:0005624; C:membrane fraction; IDA.  
 DR GO: GO:0003846; F:2-acylglycerol O-acyltransferase activity; IDA.  
 DR InterPro: IPR004299; MBOAT\_fam.  
 DR Pfam: PF03062; MBOAT, 1.  
 KW Acyltransferase; Endoplasmic reticulum; Transferase; Transmembrane.  
 FT TRANSMEM 113 133  
 FT TRANSMEM 139 159 Potential.  
 FT TRANSMEM 175 195 Potential.  
 FT TRANSMEM 198 218 Potential.

FT TRANSMEM 293 313 Potential.  
 FT TRANSMEM 343 363 Potential.  
 FT TRANSMEM 417 437 Potential.  
 FT TRANSMEM 439 459 Potential.  
 FT TRANSMEM 464 484 Potential.  
 FT ACT SITE 426 426 Potential.  
 SQ SEQUENCE 498 AA; 56790 MW; E7BDD6DDCF1EC2B CRC64;  
 Query Match 87.8%; Score 2278.5; DB 1; Length 498;  
 Best Local Similarity 85.1%; Pred. No. 1,7e-174;  
 Matches 421; Conservative 29; Mismatches 34; Indels 11; Gaps 3;  
 QY 1 MGDR--GSSRRRTGSRPSHGGGPPAAAEVRDPAAGPVGAGDAPAPAP----- 51  
 DB 1 MDDRAGAGSSRRRTGSRVSGGSGPKTEDEVDAAVSPDLAGGDAAPAPAPAHNR 60  
 QY 52 MKDGDAGVSGHWEIRCHRLQDSLFSSDGSFNSYRGILNMCVMTILSNARLFLENTLY 111  
 DB 61 DDXGRTSVGDGYMDLRCHRLQDSLFSSDGSFNSYRGILNMCVMTILSNARLFLENTLY 120  
 QY 112 GILVPIQVVSFLMDPSHWPAPCLVIANFAVAFOVEKRLAVGALTEQGLLHVAN 171  
 DB 121 GILVPIQVVSFLMDPSHWPAPCVIANSIFVAAFOLEKRLAVGALTEQGLLHVAN 180  
 QY 172 LATIICFPAAVVLVESITPVGSLTAMHTILFLKESYRDVNSWC--RRARAAASAG 229  
 DB 181 LATIICFPAAVVLVESITPVGSLTAMHTILFLKESYRDVNSWC--RRARAAASAG 240  
 QY 230 KKAASAAAPHVSYDNLTYRDLYFLPAPLYCEINPFRSPRIKRFILRIILEMFT 289  
 DB 241 KKVSGAAAOQAAVSYPDNLTYRDLVYFIPAPLYCEINPFRSPRIKRFILRIILEMFT 300  
 QY 290 QLOVGLIQOMVPTIQNSMKPKMDYSRIERLKLAVPNLILVIFPYMFHSCLVAN 349  
 DB 301 QLOVGLIQOMVPTIQNSMKPKMDYSRIERLKLAVPNLILVIFPYMFHSCLVAN 360  
 QY 350 AELMFGDREPRDMNSESVTYFWQNNNIPVHKCIRHFYKPMLRGSSKMAATGVL 409  
 DB 361 AELMFGDREPRDMNSESVTYFWQNNNIPVHKCIRHFYKPMLRGSSKMAATGVL 420  
 QY 410 ASAFHEVLYSVPLFMFLMAFTGMAQIPLAMVGRFGQNYGNAAVLSLITGPIAV 469  
 DB 421 TSAPFHEVLYSVPLFMFLMAFTGMAQIPLAMVGRFGQNYGNAAVLSLITGPIAV 480  
 QY 470 LMYVDYVLYNVEAP 484  
 DB 481 LMYVDYVLYNVEAP 495  
 Db  
 RESULT 7  
 OS4AA6 MOUSE PRELIMINARY; PRT; 498 AA.  
 ID OS4AA6; 13-SEP-2005 (TREMblrel. 31, Created)  
 AC OS4AA6; 13-SEP-2005 (TREMblrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMblrel. 31, Last annotation update)  
 DE Diacylglycerol acyltransferase.  
 GN Name=Dgat1; Synonym=Dgat;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Yamasaki Y., Watanabe T.K., Tanigami A.;  
 RT "Mus musculus diacylglycerol acyltransferase (Dgat) mRNA, complete  
 RT cds";  
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AB057816; BAC6171.1; -; mRNA.  
 DR MGI: MGI:133825; Dgat1.  
 DR GO: GO:0016021; C:Integral to membrane; TAS.  
 DR GO: GO:0005624; C:membrane fraction; IDA.







QY 209 FSYDVNSWCRBARAKASAGKASAAAPHTVSYDNLTYRDLYFLPAFLCYELNFP 268  
DB 121 FSYDVNSWCRBARAKASAGKASAAAPHTVSYDNLTYRDLYFLPAFLCYELNFP 180  
QY 269 RSPRIKRFLLRLILEMFLPTQLQVGLIQOMWVPTIQNSMKRPFKMDYSRIERLLKLAV 328  
DB 181 RSPRIKRFLLRLILEMFLPTQLQVGLIQOMWVPTIQNSMKRPFKMDYSRIERLLKLAV 240  
QY 329 PNLILWILFFWYLFHSCINAAVAELMQFGDREFRYDMWNSSESTYTPWQWNNIPVHKMCI RH 388  
DB 241 PNLILWILFFWYLFHSCINAAVAELMQFGDREFRYDMWNSSESTYTPWQWNNIPVHKMCI RH 300  
QY 389 FYKMLRGRSSKMMARCTGVPLASAFHEHYLVSPRLMFLMAFTGMAQIPLAMFVGRFF 448  
DB 301 FYKMLRGRSSKMMARCTGVPLASAFHEHYLVSPRLMFLMAFTGMAQIPLAMFVGRFF 360  
QY 449 QGNYGNAAVMLSLIIGOPIAVLMVHYDYVLYNTEAPAAEA 488  
DB 361 QGNYGNAAVMLSLIIGOPIAVLMVHYDYVLYNTEAPAAEA 400  
RESULT 11  
Q6P3J0 BRARE PRELIMINARY; PRT; 499 AA.  
AC Q6P3J0  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DE Hypothetical protein zgc:77691.  
GN ORFNames=zgc:77691;  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxID=7955;  
[1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Whole;  
RA MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Scheiner C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,  
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Whole;  
RA Strausberg R.L.;  
RA Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC063970; AAH63970.1; -; mRNA.  
DR ZFIN; ZDB-GENE-030131-4600; zgc:77691.  
DR InterPro; IPR004299; MBOAT\_fam.  
DR Pfam; PF03062; MBOAT; 1.  
KM Hypothetical protein  
SQ SEQUENCE 499 AA; 57151 MW; 3B3E3DA7685AEDF4 CRC64;

Query Match 65.0%; Score 1686; DB 2; Length 499;  
Best Local Similarity 63.8%; Pred. No. 7,4e-127;

Matches 323; Conservative 58; Mismatches 97; Indels 28; Gaps 8;  
QY 1 MGDR---GSSRRRTGSRPSHGCGPAAAEVYDAAGDVGAGAPAPAPNK-DGD 56  
DB 1 MGDRNEKGSAGHRRRTT-----SGEAYVQAARGGAEE--LSQVKEKEQKXENA 51  
QY 57 AGVSGHWE-----LRCHRLDLSFSSDGSFNYRGLIMCVMLISNARPLENLIK 111  
DB 52 AGRQKNSDAGEDPFSCHKLQESLLSSASGFSNYRGLIMCVMLVLSNARPLENLIK 111  
QY 112 GILVDPPIQVSLFLKDPHSNAPCLVLAANYEVAALQVEKRLAVGLTEQAGLLIVAN 171  
DB 112 GVLVDPPIQVSLFLKDPHSNAPCLVLAANYEVAALQVEKRLAVGLTEQAGLLIVAN 171  
QY 172 LATILCPFAAVLVLLVESTITPGSLLAIAHTILFLKFSYDVNSWCRBAR-ARASAGK 230  
DB 172 LSAILFPVAGVLSLSTVTPGVSMALSTCVLFLKLYSTYDVNSWCRBAR-ARASAGK 231  
QY 231 -----KASAAAPHTVSYDNLTYRDLYFLPAFLCYELNFPSPRIKRFLLRL 282  
DB 232 SHSCPSYKANGTAGYHTVTPGNLTHRDYVFAFAPFLCYELNFPSPRIKRFLLRL 291  
QY 283 LEMLFPTQLQVGLIQOMWVPTIQNSMKRPFKMDYSRIERLLKLAVPNHILWILFFWYLF 342  
DB 292 LEMLFPTQLQVGLIQOMWVPTIQNSMKRPFKMDYSRIERLLKLAVPNHILWILFFWYLF 351  
QY 343 HSCINAAVAELMQFGDREFRYDMWNSSESTYTPWQWNNIPVHKMCI RHFFKPMRLRGSKMM 402  
DB 352 HSMNPFVAELMQFGDREFRYDMWNSSTIPIYNSMNNIPVHKMCI RHFFKPMRLRGSKMM 411  
QY 403 ARTGVFLASAFHEHYLVSPRLMFLMAFTGMAQIPLAMFVGRFFQGNNGNAAVMLSLI 462  
DB 412 AQIAVFFLSAFHEHYLVSPRLMFLMAFTGMAQIPLAMFVGRFFQGNNGNAAVMLSLI 471  
QY 463 IGPPIAVLMVHYDYVLYNTEAPAAEA 488  
DB 472 IGPPIAVLMVHYDYVLYNTEAPAAEA 496  
RESULT 12  
Q6DHG2 BRARE PRELIMINARY; PRT; 507 AA.  
AC Q6DHG2  
DT 25-OCT-2004 (TREMBLrel. 28, Created)  
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
DE Zgc:92327.  
GN ORFNames=zgc:92327;  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxID=7955;  
[1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Whole;  
RA MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Scheiner C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,  
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Whole;  
RA Struhsberg R.;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC076012; AAH76012.1; -; mRNA.  
DR ZFIN: ZDB-GENE-040718-158; zgc:92327.  
DR InterPro: IPR004289; MBOAT\_fam.  
DR Pfam: PF03062; MBOAT; 1.  
SQ SEQUENCE 507 AA; 58757 MW; AA481F09E067FA9 CRC64;

Query Match 63.6%; Score 1648.5; DB 2; Length 507;  
Best Local Similarity 64.7%; Pred. No. 7.8e-124;  
Matches 306; Conservative 58; Mismatches 72; Indels 37; Gaps 4;

QY 41 GAAADAPAPAPKDDAGVSGHWM-----ELRCRLQDPLFS 77  
DB 32 GAAATPAA-----ADSGKSGHKARSLAKNLQNEKRCRDYDRSGCKLQESMIS 85  
QY 78 SDGFSNTRGILNMCVMIISNARILENLKYGILVDPQIVSILFKDPSWAPCLV 137  
DB 86 SASFSKYNRGILNMCVMIISNARILENLKYGILVDPQIVSILFKDPSWAPCLV 145  
QY 138 IAAANPAVAAPVEKRLANGALTEQAGILLHVAANLATICFPAAVVLVESTTPGSLA 197  
DB 146 IVCNPFILVALYTERKTLANGSISEKGLIYIFNLITLIFCPMVVVLKLPSTIPVGAPA 205  
QY 198 IMAHTILFLKLFPSYDVNSWCR-RARAKASAGKXSAAPHT-----VSPDNLTY 249  
DB 206 LGIYITILFLKXSYDVNWKCKERTQAKRSLSRSLSCSPSTSSMSQSYSTGKNTL 265  
QY 250 RDLYVFLPAPYLCELYNPPSPRIKRFILRLILEMLFPTQLQVGLIQMVPPTIQNSMK 309  
DB 266 RDLYVFLPAPYLCELYNPPSPRIKRFILRLILEMLFPTQLQVGLIQMVPPTIQNSMK 325  
QY 310 PKKMDYGRITIERLKLANPNHLIWLIFPYMLFHSCLNAVAELMOGDRFPRDWMNSG 369  
DB 326 PLOEDYTRMTBRLLRLAVPNHLIWLIFPYSPFSNMFMALRLRGDRFPRDWMNSG 385  
QY 370 VTYFQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVSPLMFRIM 429  
DB 386 IITYFQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVSPLMFRIM 445  
QY 430 AFTGMAOIPLAMPVGRFFQNGYNAAVWLSLIIIGQPIAVLMYVDYVLYL 482  
DB 446 AFMGMAOIPLAMPVGRFFQNGYNAAVWLSLIIIGQPIAVLMYVDYVLYL 498

## RESULT 13

Q4RLB7.TETNG  
Q4RLB7.TETNG PRELIMINARY; PRT; 473 AA.

DT 13-SEP-2005 (TREMBlrel. 31, Created)  
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
DE Chromosome 21 SCAF15022, whole genome shotgun sequence.  
(Fragment).

OS ORFNames=GSTENG00032560001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorphi; Acanthopterygii; Perciformes; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCBI\_TaxID=99883;  
RN [1]

## NUCLEOTIDE SEQUENCE.

RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
RA Mauceli E., Bounneau L., Fischer C., Ozout-Costaz C., Bernot A.,  
RA Niclaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
RA Anthouard V., Jubin C., Castellini V., Katinin M., Vacherie B.,  
RA Bismont C., Skalli Z., Catalloico L., Boulain J., De Berardinis V.,  
RA Cruaud C., Duprat S., Brottier P., Couranceau J.P., Gouzy J.,  
RA Parra G., Lardier G., Chappe C., McKernan K.J., McEwan P., Bosak S.,  
RA Kellis M., Voigt J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Landet V., Schachter V., Queller F., Saurin W., Scarpelli C.,  
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.;  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
the early vertebrate proto-karyotype";  
RL Nature 431:946-957 (2004).

RP NUCLEOTIDE SEQUENCE.  
RG Genoscope; Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL: CAAB01015022; CAG10815.1; -; Genomic\_DNA.  
FT NON TER 1  
FT NON TER 1  
SQ SEQUENCE 473 AA; 53366 MW; 0CE0F20F5093342 CRC64;

Query Match 62.6%; Score 1624; DB 2; Length 473;  
Best Local Similarity 63.8%; Pred. No. 6.7e-122;  
Matches 301; Conservative 51; Mismatches 62; Indels 58; Gaps 3;

QY 68 CHRLQDLSFSDSGFSNTRGILNMCVMI-----ILSNA 101  
DB 1 CHVLOESLSSASGSNTRGILNMCVMIILFGGFYASGALGHQGLLIDAFPMKAVLSNA 60  
QY 102 RLEENLTKYGILVDPQIVSILFKDPSWAPCLVIANVAVAFOVEKRLANGALTE 161  
DB 61 RLEENLTKYGILVDPQIVSILFKDPSWAPCLVIANVAVAFOVEKRLANGALTE 120  
QY 162 QAGLLHVAANLATICFPAAVVLVESTTPGSLALMAHTILFLKLFPSYDVNSWCRRA 221  
DB 121 TTRGILHFNLASLILFPSATVLTVPSTIPVGVSISLGIYTLFLKXSYGQTNMCRRI 180  
QY 222 R-AKASAGKXSAAPHTVSYDNLTYRDLTYFLPAPYLCELYNPPSPRIKRFILR 280  
DB 181 R-AKASAGKXSAAPHTVSYDNLTYRDLTYFLPAPYLCELYNPPSPRIKRFILR 240  
QY 281 RILEM-----LEFQQLQVGLIQMVPPTIQNSMK 309  
DB 241 RLEFVIRGNAYAAARYSILIPNRRITLFLVHFCILQLEFFMQLVGLIQMVPPTIQNSMK 300  
QY 310 PKKMDYGRITIERLKLANPNHLIWLIFPYMLFHSCLNAVAELMOGDRFPRDWMNSG 369  
DB 301 PLOEDYTRMTBRLLRLAVPNHLIWLIFPYSPFSNMFMALRLRGDRFPRDWMNSG 360  
QY 370 VTYFQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVSPLMFRIM 429  
DB 361 VTYFQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVSPLMFRIM 420  
QY 430 AFTGMAOIPLAMPVGRFFQNGYNAAVWLSLIIIGQPIAVLMYVDYVLYL 481  
DB 421 AFMGMAOIPLAMPVGRFFQNGYNAAVWLSLIIIGQPIAVLMYVDYVLYL 472

## RESULT 14

Q4RI38.TETNG  
Q4RI38.TETNG PRELIMINARY; PRT; 447 AA.

DT 13-SEP-2005 (TREMBlrel. 31, Created)  
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
DE Chromosome 8 SCAF15044, whole genome shotgun sequence.  
(Fragment).

OS ORFNames=GSTENG00034052001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 NCBI\_taxid=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomas N.,  
 Maucell E., Bouneau L., Fischer C., Orouf-Costas C., Bernot A.,  
 Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 Daesli C., Salenoubat M., Levy M., Boudet N., Castellano S.,  
 Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 Bismont C., Skalli Z., Caticolico L., Poulain J., De Berardinis V.,  
 Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,  
 Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,  
 Kellis M., Volff J.N., Guitto R., Zody M.C., Mesirov J.,  
 Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 Landet V., Schachter V., Queller F., Sautin W., Scarpelli C.,  
 Winkler P., Lander E.S., Weissbach J., Roest Crolians H.,  
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 the early vertebrate proto-karyotype";  
 RL Nature 431:946-957(2004).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.  
 DR EMBL; CAAB01015044; CAG11944.1; -; Genomic\_DNA.  
 FT NON\_TER 1 1  
 FT SEQUENCE 447 AA; 52106 MW; E335A2785B3CDBFA CRC64;  
 SQ  
 Query Match 62.0%; Score 1608; DB 2; Length 447;  
 Best Local Similarity 66.4%; Pred. No. 1.2e-120;  
 Matches 296; Conservative 52; Mismatches 70; Indels 28; Gaps 4;  
 QY 68 CHR.LDDSLPSSDGSFNSYNGILNMCVVMILISNARLFLENLKYGLVDPIDVNSFLKD 127  
 DB 1 CHKNDLSLSASGSKRGILNMCVVMILISNARLFLENLKYGLVDPIDVNSFLND 60  
 QY 128 PHSWAPCLVIAANYFAVAFAOVERKRLAVGALTQAGLLLVANLTLICPPAAVLLVE 187  
 DB 61 PYSWPAACLVIAANYFAVAFAOVERKRLAVGALTQAGLLLVANLTLICPPAAVLLVE 120  
 QY 188 SITP-----VGSLLALMAHTLFLKLFSDYRDVNSMCR-ARAKASGKXASGAANH 239  
 DB 121 SVTPSVQSSGVGAASALSITLFLKLYSKDVNLWCRESLTIKVKLSRLSCPSQOH 180  
 QY 240 -----TVSYPDNLTYRDLYFLFAPTLCYELNPPSPRIRKFLRLILEMLFTLOLY 293  
 DB 181 PSGGCKKSYSPGNLTLRDNYTFVPAFTLCYELNPPSPRIRKFLRLILEMLFTLOLY 240  
 QY 294 GLIQMNVPTIONSMPKMDYSRIIERLLKLAVERNHLIWLIPYWLFSCLNAVAELM 353  
 DB 241 ALTQGMVPIIRSSMKPLEDMDSRAVERLLRLAVNHLMLMFYWFPHSSMNTAEEL 300  
 QY 354 QPDGDEPRDWNSSVYTFPQNMNIPVHKMCIRHFYKMLRSGSKXMAATGCVPLASAF 413  
 DB 301 RFGDRQFYNDWNSEVYTFPQNMNIPVHKMCIRHFYKMLRSGSKXMAATGCVPLASAF 360  
 QY 414 PHBYLVSVPLRMFLMAFTGMAA-----QIPLAFVGRFQNGYGNAAVWL 459  
 DB 361 PHBYLVSVPLRMFLMAFTGMAA-----QIPLAFVGRFQNGYGNAAVWL 420  
 QY 460 SLIIQPIAVLMLYVHDYVLYVLAEP 485  
 DB 421 SLIIQPIAVLMLYVHDYVLYVLAEP 446

RESULT 15  
 Q91YB5\_RAT PRELIMINARY; PRT; 242 AA.

AC Q91YB5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Diacylglycerol acyltransferase I (EC 2.3.1.20) (Fragment).  
 GN Name=dgat1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Rattus.  
 NCBI\_taxid=10116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Mistar; TISSUE=Liver;  
 RA Waterman J.J., Zammit V.A., Price N.T.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ345014; CAC69884.1; -; mRNA.  
 DR GO; GO:0008415; F:acyltransferase activity; IEA.  
 DR GO; GO:0004144; F:diacylglycerol O-acyltransferase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR004299; MBOAT\_fam.  
 DR Pfam; PF03062; MBOAT; 1  
 KW Acyltransferase; Transferase.  
 FT NON\_TER 1 1  
 FT SEQUENCE 242 AA; 29014 MW; E12F6DA30988542 CRC64;  
 SQ  
 Query Match 47.4%; Score 1230; DB 2; Length 242;  
 Best Local Similarity 92.5%; Pred. No. 1.4e-90;  
 Matches 221; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
 QY 219 RRARAKASAGKKAASAAAPHTVSYPDNLTYRDLYFLFAPTLCYELNPPSPRIRKFL 278  
 DB 4 RRVAKASAGKKAASAAAPHTVSYPDNLTYRDLYFLFAPTLCYELNPPSPRIRKFL 63  
 QY 279 LRRILEMLFTLOLYGLIQMNVPTIONSMPKMDYSRIIERLLKLAVERNHLIWLIF 338  
 DB 64 LRRILEMLFTLOLYGLIQMNVPTIONSMPKMDYSRIIERLLKLAVERNHLIWLIF 123  
 QY 339 YWLPHSCLNAVAELMOPGDRFVYDWNSSVYTFPQNMNIPVHKMCIRHFYKMLRSGS 398  
 DB 124 YWLPHSCLNAVAELMOPGDRFVYDWNSSVYTFPQNMNIPVHKMCIRHFYKMLRSGS 183  
 QY 399 SKWMARTGVPLASAFHFHYLVSVPLRMFLMAFTGMAAQIPLAFVGRFQNGYGNAAV 457  
 DB 184 SKWMARTGVPLASAFHFHYLVSVPLRMFLMAFTGMAAQIPLAFVGRFQNGYGNAAV 242

Search completed: May 5, 2006, 12:33:32  
 Job time : 234 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

## OM protein - protein search, using sw model

Run on: May 5, 2006, 12:27:56 / Search time 88 Seconds  
(without alignments)  
2436.555 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MDRGSSRRRTGRSPSSHG.....VLMYVDYVLYNPAAPAEA 488

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

. Listing first 45 summaries

Database : A\_Geneseq\_21:.\*  
1: geneseqp19808:.\*  
2: geneseqp19908:.\*  
3: geneseqp20008:.\*  
4: geneseqp20018:.\*  
5: geneseqp20028:.\*  
6: geneseqp20038:.\*  
7: geneseqp20039s:.\*  
8: geneseqp20048:.\*  
9: geneseqp20058:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query Length	ID	Description
1	2594	100.0	488	3	AA015200 Human ACA
2	2594	100.0	488	8	AD000389 Human dia
3	2594	100.0	488	8	AD084845 Human dia
4	2588	99.8	488	7	ADD45554 Human Pro
5	2588	99.8	488	7	AD059053 Human Pro
6	2588	99.8	488	8	ABW81016 Tumour-as
7	2588	99.8	488	8	ADQ39678 Human myo
8	2588	99.8	488	8	ADQ39678 Human myo
9	2293.5	88.6	500	3	AA054139 Acyl-CoA:
10	2293.5	88.6	500	3	AA054139 Acyl-CoA:
11	2292.5	88.4	489	5	AA024973 Bovine DG
12	2286.5	88.1	489	5	AA024973 Bovine di
13	2278.5	87.8	489	6	ADP96046 Bovine DG
14	2278.5	87.8	489	6	AD016606 Mouse dig
15	2278.5	87.8	489	8	AD000390 Mouse dia
16	2278.5	87.8	489	8	AD084846 Mouse dia
17	2278.5	87.6	498	8	AA045622 Mouse dia
18	2273.5	87.6	498	6	ADP96045 Bovine DG
19	2266.5	87.4	498	7	AD059051 Rat Prote
20	2266.5	87.4	498	7	AD045552 Rat Prote
21	2266.5	87.4	498	8	AD084847 Rat diacy
22	2264	87.3	497	3	AA094523 Bovine dia
23	2192.5	84.5	467	5	AA024974 Bovine di
24	2148	82.8	407	2	AA043406 Human acy

25	2046	78.9	387	8	ADQ39677 Human myo
26	2039	78.6	386	3	AA045621 Human dia
27	1681	64.8	421	8	AD035335 ENZM prot
28	1681	64.8	421	8	ADQ39679 Human myo
29	1297	50.0	236	4	AB010191 Human cdn
30	1297	50.0	236	5	ADP66778 Human pol
31	1159	44.7	219	2	AA043412 Human acy
32	1106	42.6	496	3	AA054140 Acyl-CoA:
33	1106	42.6	496	3	AA019741 Caenorhab
34	1106	42.6	496	8	AD023832 Bacterial
35	1106	42.6	496	8	AD023831 Bacterial
36	834	32.2	538	8	AD080111 Plant ful
37	832.5	32.1	496	8	AD011508 Soybean d
38	819	31.6	507	9	AD016878 Eucorymus
39	819	31.6	507	9	AD016893 Eucorymus
40	807	31.1	508	3	AA094522 Wheat dia
41	800.5	30.9	504	3	AA094519 Soybean d
42	798.5	30.8	520	3	AA094512 Arabidops
43	797	30.7	532	8	AD081158 Tobacco d
44	797	30.7	532	8	AD081145 Tobacco d
45	791.5	30.5	520	3	AA054143 Acyl-CoA:

## ALIGNMENTS

RESULT 1	AA015200	standard, protein, 488 AA.
ID	AA015200	standard, protein, 488 AA.
AC	AA015200;	
XX		
DT	14-DEC-2000	(first entry)
DE	Human ACAT Related Gene Product 1 ARGP1.	
KW	Human; ACAT Related Gene Product 1; ARGP1; gene therapy; enzyme;	
KW	acyl Coenzyme A-cholesterol acyltransferase 1; ACAT1;	
KW	sterol esterification; lipid homeostasis; diacylglycerol acyltransferase;	
KW	DGAT; cholesterol; triglyceride biosynthesis; hypertriglyceridaemia;	
KW	hyperlipidaemia; atherosclerosis; heart disease; obesity.	
OS	Homo sapiens.	
XX		
PN	US6100077-A.	
PD	08-AUG-2000.	
XX		
PF	01-OCT-1998;	98US-00165042.
XX		
PR	01-OCT-1998;	98US-00165042.
XX		
PA	(UYCO ) UNIV COLUMBIA NEW YORK.	
XX		
PI	Sturley SL, Oelkers P;	
XX		
DR	WPI: 2000-557622/51.	
DR	N-PSDB; AAA76169.	
XX		
PT	New nucleic acid encoding a human diacylglycerol acyltransferase, useful	
PT	for treating hyperlipidaemia, atherosclerosis, heart disease, or other	
PT	diseases associated with an imbalance of triglyceride levels.	
XX		
PS	Claim 1; Fig 1A; 32pp; English.	
XX		
CC	The enzyme acyl Coenzyme A-cholesterol acyltransferase 1 (ACAT1) mediates	
CC	sterol esterification, an important component of intracellular lipid	
CC	homeostasis. The present sequence is human ACAT Related Gene Product 1	
CC	(ARGP1). This enzyme is a diacylglycerol acyltransferase (DGAT). This	
CC	enzyme does not esterify cholesterol. It is thought therefore that ARGP1	
CC	participates in the Coenzyme A-dependent acylation of substrate(s) other	
CC	than cholesterol e.g. diacylglycerol. Also, ARGP1 has a predicted	
CC	diacylglycerol binding motif, suggesting that it may perform the last	

CC acylation in triglyceride biosynthesis. ACP1 gene and protein are useful  
CC for treating a subject who has an imbalance in triglyceride levels due to  
CC a defect in esterification of diacylglycerol via gene therapy. Particularly,  
CC ACP1 is useful for treating hypertriglyceridaemia, hyperlipidaemia,  
CC atherosclerosis, heart disease, obesity or other diseases associated with  
CC high or excessive levels of triglyceride

XX Sequence 488 AA:

Query Match 100.0%; Score 2594; DB 3; Length 488;  
Best Local Similarity 100.0%; Pred. No. 2.4e-273;  
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDSSRRRRRGTGSPSSHGCGPAAAEYRDAAAGPDVGAAGDAPAPAPNKKDAGVG 60  
DB 1 MGDSSRRRRRGTGSPSSHGCGPAAAEYRDAAAGPDVGAAGDAPAPNKKDAGVG 60  
QY 61 SGHWELRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKYGILVDP1QV 120  
DB 61 SGHWELRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKYGILVDP1QV 120  
QY 121 VSLFKDPSHPAPCLVIAANFAVAAFQVEKRLAVGALTQAGLLHVNALATILCPA 180  
DB 121 VSLFKDPSHPAPCLVIAANFAVAAFQVEKRLAVGALTQAGLLHVNALATILCPA 180  
QY 181 AVVLIVESITPVGSLALMAHTILFLKLFYRDVNSWCRARAKAASAGKASAAAPHT 240  
DB 181 AVVLIVESITPVGSLALMAHTILFLKLFYRDVNSWCRARAKAASAGKASAAAPHT 240  
QY 241 VSYPDNLTYRDLYYFLFAPTLCELNFPSPRIIRKRFLLRLILEMFFTOLOVGILQOMM 300  
DB 241 VSYPDNLTYRDLYYFLFAPTLCELNFPSPRIIRKRFLLRLILEMFFTOLOVGILQOMM 300  
QY 301 VPTIIONSKPKPKMDYSRIIERLLKLAVPNHLIMLFFYWLPHSCINAAVAELMOGDBEF 360  
DB 301 VPTIIONSKPKPKMDYSRIIERLLKLAVPNHLIMLFFYWLPHSCINAAVAELMOGDBEF 360  
QY 361 YRDWNSSESVTFMQWNMIPVHKCIRHFYKPMRLRGSSKMMARGTGVLASAFHEIYVS 420  
DB 361 YRDWNSSESVTFMQWNMIPVHKCIRHFYKPMRLRGSSKMMARGTGVLASAFHEIYVS 420  
QY 421 VPLRMFLMAFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQPIAVLAVYVHDYYVLN 480  
DB 421 VPLRMFLMAFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQPIAVLAVYVHDYYVLN 480  
QY 481 YEAPAAEA 488  
DB 481 YEAPAAEA 488

RESULT 2

ADS00389 ID ADS00389 standard; protein; 488 AA.

XX ADS00389;

XX 16-DEC-2004 (first entry)

XX Human diacylglycerol acyltransferase 1.

XX Antifateriosclerotic; Antilipemic; Antidiabetic; Anorectic; Cardiant;  
XX diacylglycerol acyltransferase-inhibitor-1;  
XX diacylglycerol acyltransferase 1; glucose; cholesterol; triglyceride;  
XX lipid metabolism; cholesterol metabolism; atherosclerosis;  
XX hyperlipidaemia; diabetes; type 2 diabetes; obesity;  
XX cardiovascular disease; human; enzyme.

OS Homo sapiens.

XX US2004185559-A1.

XX 23-SEP-2004.

XX

PF 21-MAR-2003; 2003US-00394808.

XX 21-MAR-2003; 2003US-00394808.

PA (ISIS-) ISIS PHARM INC.

PI Monta BP, Graham MJ;

DR WPI: 2004-667979/65.

DR N-PSDB; ADS00240.

PT New antisense oligonucleotides targeted to diacylglycerol acyltransferase

PT 1, useful for treating diacylglycerol acyltransferase 1-associated

PT diseases or conditions, e.g. atherosclerosis, hyperlipidemia, diabetes or

PT obesity.

PS Disclosure; Page 27-29; 55pp; English.

CC The invention relates to a compound targeted to a nucleic acid molecule  
CC encoding diacylglycerol acyltransferase 1. The compound is useful for  
CC treating a disease or condition associated with diacylglycerol  
CC acyltransferase 1; modulating glucose or cholesterol levels, or lowering  
CC triglyceride levels in an animal; or for delaying the onset of a disease  
CC or condition associated with diacylglycerol acyltransferase 1 in an  
CC animal, such as abnormal lipid metabolism, abnormal cholesterol  
CC metabolism, atherosclerosis, abnormal metabolic condition,  
CC hyperlipidaemia, diabetes, type 2 diabetes, obesity, cardiovascular  
CC disease. The present sequence represents the amino acid sequence of human  
CC diacylglycerol acyltransferase 1.

XX Sequence 488 AA:

Query Match 100.0%; Score 2594; DB 8; Length 488;  
Best Local Similarity 100.0%; Pred. No. 2.4e-273;  
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDSSRRRRRGTGSPSSHGCGPAAAEYRDAAAGPDVGAAGDAPAPAPNKKDAGVG 60  
DB 1 MGDSSRRRRRGTGSPSSHGCGPAAAEYRDAAAGPDVGAAGDAPAPAPNKKDAGVG 60  
QY 61 SGHWELRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKYGILVDP1QV 120  
DB 61 SGHWELRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKYGILVDP1QV 120  
QY 121 VSLFKDPSHPAPCLVIAANFAVAAFQVEKRLAVGALTQAGLLHVNALATILCPA 180  
DB 121 VSLFKDPSHPAPCLVIAANFAVAAFQVEKRLAVGALTQAGLLHVNALATILCPA 180  
QY 181 AVVLIVESITPVGSLALMAHTILFLKLFYRDVNSWCRARAKAASAGKASAAAPHT 240  
DB 181 AVVLIVESITPVGSLALMAHTILFLKLFYRDVNSWCRARAKAASAGKASAAAPHT 240  
QY 241 VSYPDNLTYRDLYYFLFAPTLCELNFPSPRIIRKRFLLRLILEMFFTOLOVGILQOMM 300  
DB 241 VSYPDNLTYRDLYYFLFAPTLCELNFPSPRIIRKRFLLRLILEMFFTOLOVGILQOMM 300  
QY 301 VPTIIONSKPKPKMDYSRIIERLLKLAVPNHLIMLFFYWLPHSCINAAVAELMOGDBEF 360  
DB 301 VPTIIONSKPKPKMDYSRIIERLLKLAVPNHLIMLFFYWLPHSCINAAVAELMOGDBEF 360  
QY 361 YRDWNSSESVTFMQWNMIPVHKCIRHFYKPMRLRGSSKMMARGTGVLASAFHEIYVS 420  
DB 361 YRDWNSSESVTFMQWNMIPVHKCIRHFYKPMRLRGSSKMMARGTGVLASAFHEIYVS 420  
QY 421 VPLRMFLMAFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQPIAVLAVYVHDYYVLN 480  
DB 421 VPLRMFLMAFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQPIAVLAVYVHDYYVLN 480  
QY 481 YEAPAAEA 488  
DB 481 YEAPAAEA 488

RESULT 3  
 ADS84845  
 ID ADS84845 standard; protein; 488 AA.  
 XX  
 AC ADS84845;  
 XX  
 DT 13-JAN-2005 (first entry)  
 XX  
 DE Human diacylglycerol acyltransferase 1 (DGAT1).  
 XX  
 KM antilipemic; antiarteriosclerotic; antidiabetic; anorectic;  
 KM cardiovascular; gene therapy; DGAT-Modulator-1; DGAT-Inhibitor-1;  
 KM diacylglycerol acyltransferase 1; DGAT1; gene expression;  
 KM DGAT1 modulator; cholesterol; glucose level; insulin level;  
 KM glucose tolerance; insulin tolerance; lipid metabolism;  
 KM cholesterol metabolism; atherosclerosis; abnormal metabolic condition;  
 KM hyperlipidemia; diabetes; Type 2 diabetes; obesity;  
 KM cardiovascular disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004209838-A1.  
 XX  
 PD 21-OCT-2004.  
 XX  
 PF 18-MAR-2004; 2004US-00803482.  
 XX  
 PR 21-MAR-2003; 2003US-00394808.  
 XX  
 PA (MONI/) MONIA B P.  
 PA (GRAH/) GRAHAM M J.  
 XX  
 PI Monica BP, Graham MJ;  
 XX  
 DR WPI: 2004-747287/73.  
 DR N-PsDB; ADS84617.  
 XX  
 PT New compound 8-80 nucleobases in length targeted to a nucleic acid  
 PT molecule encoding diacylglycerol acyltransferase 1 (DGAT1), useful for  
 PT treating an animal having e.g., atherosclerosis, hyperlipidemia,  
 PT diabetes, or obesity.  
 PT  
 PS Disclosure; Page 34-36; 77pp; English.  
 XX  
 CC The invention describes a compound 8 to 80 nucleobases in length targeted  
 CC to a nucleic acid molecule encoding diacylglycerol acyltransferase 1  
 CC (DGAT1), where the compound specifically hybridizes with the nucleic acid  
 CC molecule encoding DGAT1 comprising a sequence comprising 1976 bp (SEQ ID  
 CC NO: 4) fully defined in the specification, and inhibits the expression of  
 CC DGAT1. Also described are: inhibiting the expression of DGAT1 in cells or  
 CC tissues; screening for a modulator of DGAT1; a diagnostic method for  
 CC identifying a disease state; a kit or assay device comprising the  
 CC compound above; modulating glucose levels in an animal; preventing or  
 CC delaying the onset of a disease or condition associated with DGAT1 in an  
 CC animal; modulating cholesterol levels in an animal; reducing serum  
 CC glucose levels; DGAT1 levels in the liver; circulating insulin levels;  
 CC decreasing fasting serum insulin levels; improving an animal's performance  
 CC on glucose tolerance tests and insulin tolerance tests; reducing  
 CC circulating triglycerides, liver triglycerides or free fatty acids in the  
 CC liver of an animal; and lowering triglyceride levels in an animal. The  
 CC compound is useful in a method for treating an animal having a disease or  
 CC condition associated with DGAT1, i.e., abnormal lipid metabolism,  
 CC abnormal cholesterol metabolism, atherosclerosis, abnormal metabolic  
 CC condition (i.e., hyperlipidemia), diabetes (i.e., Type 2 diabetes),  
 CC obesity, or cardiovascular disease, so that expression of DGAT1 is  
 CC inhibited. This is the amino acid sequence of human diacylglycerol  
 CC acyltransferase 1 (DGAT1).  
 XX  
 SQ Sequence 488 AA;

QY 1 MGDGSSRRRTGSRPSSHGCGGPAAAEEVEVDAAAGPDVGAAGAPAPAPKDGDAVG 60  
 DB 1 MGDGSSRRRTGSRPSSHGCGGPAAAEEVEVDAAAGPDVGAAGAPAPAPKDGDAVG 60  
 QY 61 SGHWEI RCHRLQDSLFFSSDSCGFSNYRGILMVCVMLISNARLPLENLIKYGI LVDPIQV 120  
 DB 61 SGHWEI RCHRLQDSLFFSSDSCGFSNYRGILMVCVMLISNARLPLENLIKYGI LVDPIQV 120  
 QY 121 VSLFLKDPHSPWAPCLVIAANVAFAVQVEKRLAVGALTDEQAGILLHVAANLATILCFPA 180  
 DB 121 VSLFLKDPHSPWAPCLVIAANVAFAVQVEKRLAVGALTDEQAGILLHVAANLATILCFPA 180  
 QY 181 AVVLVVESTTPGSLTAAHTILFLKLFPSYRDVSWCERBARAKAASAGKASAAAPRT 240  
 DB 181 AVVLVVESTTPGSLTAAHTILFLKLFPSYRDVSWCERBARAKAASAGKASAAAPRT 240  
 QY 241 VSPYDNLTYRDLVYFLFAPTLCYELNPPRSPIRRKRLIRILEMLFFTQIQVGLIIQWM 300  
 DB 241 VSPYDNLTYRDLVYFLFAPTLCYELNPPRSPIRRKRLIRILEMLFFTQIQVGLIIQWM 300  
 QY 301 VPTIIONSKKPKDMYSRILIERLLGLAVPNHLIWLIFPYWLPHSCLNVAELMGGDEEF 360  
 DB 301 VPTIIONSKKPKDMYSRILIERLLGLAVPNHLIWLIFPYWLPHSCLNVAELMGGDEEF 360  
 QY 361 YRDMNNSSVTYFMQNMNI PVHKICIRPFYKPMLERRGSKMMARGVFLAASFPEHYLVS 420  
 DB 361 YRDMNNSSVTYFMQNMNI PVHKICIRPFYKPMLERRGSKMMARGVFLAASFPEHYLVS 420  
 QY 421 VPLRMFRLMAFTGMAAQIPLAMFVGRFPGYGNAAVWLSLIIGPIAVLWVHDYYVLN 480  
 DB 421 VPLRMFRLMAFTGMAAQIPLAMFVGRFPGYGNAAVWLSLIIGPIAVLWVHDYYVLN 480  
 QY 481 YEAPAAEA 488  
 DB 481 YEAPAAEA 488

RESULT 4  
 ADD45554  
 ID ADD45554 standard; protein; 488 AA.  
 XX  
 AC ADD45554;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein XP\_035370. SEQ ID NO 11218.  
 XX  
 KM Human; pain; neuronal tissue; gene therapy;  
 KM spinal segmental nerve injury; chronic constriction injury; CCI;  
 KM spared nerve injury; SNJ; Chung.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GHEO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 DR WPI: 2003-268312/26.  
 DR GENBANK; XP\_035370.  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page: 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 488 AA;

XX Query Match 99.8%; Score 2588; DB 7; Length 488;

XX Best Local Similarity 99.8%; Pred. No. 1.1e-272;

XX Matches 487; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRTGSPSSHGGGPAABEEVDAAAGPVGAGAPAPAPKDDAGVG 60  
DB 1 MGDGSSRRRTGSPSSHGGGPAABEEVDAAAGPVGAGAPAPAPKDDAGVG 60  
QY 61 SGHWELRCHRLDLSFSSDSGFSNYRGILMVCVWLILSNARLFLENLTKYGIIVDP1QV 120  
DB 61 SGHWELRCHRLDLSFSSDSGFSNYRGILMVCVWLILSNARLFLENLTKYGIIVDP1QV 120  
QY 121 VSLFLKDHSPAPCLVIAANFANAQVEKRLAVGALTQAGLLHVAANLATTLCPPA 180  
DB 121 VSLFLKDHSPAPCLVIAANFANAQVEKRLAVGALTQAGLLHVAANLATTLCPPA 180  
QY 121 VSLFLKDEYSWPAFLVIAANFAVAAPQVEKRLAVGALTQAGLLHVAANLATTLCPPA 180  
DB 121 VSLFLKDEYSWPAFLVIAANFAVAAPQVEKRLAVGALTQAGLLHVAANLATTLCPPA 180  
QY 181 AVLLVVESTIPVGSLLALMAHTILFLKFSYRDVNSWCRBARAKASAGKSSAAAPHT 240  
DB 181 AVLLVVESTIPVGSLLALMAHTILFLKFSYRDVNSWCRBARAKASAGKSSAAAPHT 240  
QY 241 VSYPDNLTYRDLTYFLFAPPTLCYELNFPSPRIKRRFLRLILEMLFTQIQVGLIQWM 300  
DB 241 VSYPDNLTYRDLTYFLFAPPTLCYELNFPSPRIKRRFLRLILEMLFTQIQVGLIQWM 300  
QY 301 VPTIIONSKKPPKDMYSRILIERLLKLANVNHILWILFFYVWLFHSCLANVAELMGGDREF 360  
DB 301 VPTIIONSKKPPKDMYSRILIERLLKLANVNHILWILFFYVWLFHSCLANVAELMGGDREF 360  
QY 361 YRDWNSNSVITYFQONMNI PVHKWCIRHFYKPMLRGSSKMAARGVFLASAFPEHYVS 420  
DB 361 YRDWNSNSVITYFQONMNI PVHKWCIRHFYKPMLRGSSKMAARGVFLASAFPEHYVS 420  
QY 421 VPLRMFLMAFTGMAAQIPLANFVGRFQNGYGNAAWLSLIIIGPIAVLMTVHDYYVLN 480  
DB 421 VPLRMFLMAFTGMAAQIPLANFVGRFQNGYGNAAWLSLIIIGPIAVLMTVHDYYVLN 480  
QY 481 YEAPAAEA 488  
DB 481 YEAPAAEA 488

RESUT 5  
ADE59053  
ID ADE59053 standard; protein: 488 AA.

AC ADE59053;

DT 29-JAN-2004 (first entry)

DE Human Protein XP\_035370, SEQ ID NO 4944.

KM Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KX spared nerve injury; SNI; Chung.

OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-266312/26.

DR GENBANK; XP\_035370.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page: 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 488 AA;

XX Query Match 99.8%; Score 2588; DB 7; Length 488;

XX Best Local Similarity 99.8%; Pred. No. 1.1e-272;





AD039676  
ID AD039676 standard; protein; 488 AA.  
XX  
AC AD039676;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human myocardial infarction-associated gene derived protein, SEQ ID 1339.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2004058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
PA (APPL-) APPLERA CORP.  
XX  
PI Cargill M, Devlin JT, Iakubova O;  
XX  
DR WPI; 2004-533949/51.  
DR N-PSDB; AD038848.  
XX  
PT Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX  
PS Claim 10; SEQ ID NO 1339; 145pp; English.  
XX  
CC The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiant activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC sequence represents the protein of a human myocardial infarction-  
CC associated gene containing one or more SNP's of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.  
XX  
SQ Sequence 488 AA;

Query Match 99.8%; Score 2588; DB 8; Length 488;  
Best Local Similarity 99.8%; Pred. No. 1.1e-272;  
Matches 487; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 MGDGSSRRRTGSRPSSHGCGPAAABEEVDAAGDPVGAAGDAPAPAPNKGDAGVG 60

|||||  
Db 1 MGDGSSRRRTGSRPSSHGCGPAAABEEVDAAGDPVGAAGDAPAPAPNKGDAGVG 60  
QY 61 SGHWEIRCHRIQDSLFSSSDSGFSNVRGILNMCVMYLLISNARLFLENLIKXGIIIVDPYQV 120  
Db 61 SGHWEIRCHRIQDSLFSSSDSGFSNVRGILNMCVMYLLISNARLFLENLIKXGIIIVDPYQV 120  
QY 121 VSLFLKDPHSWPAPCLVYAAVFAVFAAFQVEKRLAVGALTQAGLLHVNALATILCPFA 180  
Db 121 VSLFLKDPHSWPAPCLVYAAVFAVFAAFQVEKRLAVGALTQAGLLHVNALATILCPFA 180  
QY 181 AVVLIVESITPVGSLALMAHTILFLKFSYRDVNSWCRARAPAKAASAGKKAASSAAAPHT 240  
Db 181 AVVLIVESITPVGSLALMAHTILFLKFSYRDVNSWCRARAPAKAASAGKKAASSAAAPHT 240  
QY 241 VSYPDNLTYRDLTYFLFAPTCYELNFPERSRIKRFLLRILLMFLPTQLOVGLIOQWM 300  
Db 241 VSYPDNLTYRDLTYFLFAPTCYELNFPERSRIKRFLLRILLMFLPTQLOVGLIOQWM 300  
QY 301 VPTIONSMKPFKMDYSRIIRLLKLAVPNHLIMLIFPYWLFHSCLNVAELMFGDREF 360  
Db 301 VPTIONSMKPFKMDYSRIIRLLKLAVPNHLIMLIFPYWLFHSCLNVAELMFGDREF 360  
QY 361 YRDMWNSSESVTFYQNNWNI PVHKWCIRHFYFPMRLRGSSKMMARTGFLASAFHEHYLV 420  
Db 361 YRDMWNSSESVTFYQNNWNI PVHKWCIRHFYFPMRLRGSSKMMARTGFLASAFHEHYLV 420  
QY 421 VPLRMFRLMAFTGMAAOIPLAMFVGRFQGNVGAALWLSLITGPIAVLMYVDYVYLN 480  
Db 421 VPLRMFRLMAFTGMAAOIPLAMFVGRFQGNVGAALWLSLITGPIAVLMYVDYVYLN 480  
QY 481 YEAPAAEA 488  
Db 481 YEAPAAEA 488

RESULT 8  
AD039678  
ID AD039678 standard; protein; 488 AA.  
XX  
AC AD039678;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human myocardial infarction-associated gene derived protein, SEQ ID 1341.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2004058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
PA (APPL-) APPLERA CORP.  
XX  
PI Cargill M, Devlin JT, Iakubova O;  
XX  
DR WPI; 2004-533949/51.  
DR N-PSDB; AD038850.  
XX  
PT Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX

PS Claim 10; SEQ ID NO 1341; 145bp; English.

XX The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiac activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC sequence represents the protein of a human myocardial infarction-  
CC associated gene containing one or more SNP's of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.

XX Sequence 488 AA;

Query Match 99.8%; Score 2588; DB 8; Length 488;

Best Local Similarity 99.8%; Pred. No. 1.1e-272; Matches 487; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRTGSPSSSHGCGPAAAEBEVDAAAGPVGAAGAPAPAPKDGAGVG 60  
DB 1 MGDGSSRRRTGSPSSSHGCGPAAAEBEVDAAAGPVGAAGAPAPAPKDGAGVG 60  
QY 61 SGHWELRCHRLQDSLFFSSDGSFNSYRGILNMCVMLILSNARLFLENLIKYGILVDP10V 120  
DB 61 SGHWELRCHRLQDSLFFSSDGSFNSYRGILNMCVMLILSNARLFLENLIKYGILVDP10V 120  
QY 121 VSLFKDHSWAPCLVTAANFAVAARQVEKRLAVGALTBOAGLILHMANLATTICPA 180  
DB 121 VSLFKDHSWAPCLVTAANFAVAARQVEKRLAVGALTBOAGLILHMANLATTICPA 180  
QY 181 AVVLVESITPVGSLAAMAHTILFLKFSYRDVNSWCRBARAKASAKKSSAAAPHT 240  
DB 181 AVVLVESITPVGSLAAMAHTILFLKFSYRDVNSWCRBARAKASAKKSSAAAPHT 240  
QY 241 VSYDNLTYRDLTYFLFAPTLCYELNPPSPRIKRRFLIRLILEMLFTQLQVGIQOM 300  
DB 241 VSYDNLTYRDLTYFLFAPTLCYELNPPSPRIKRRFLIRLILEMLFTQLQVGIQOM 300  
QY 301 VPTIONSKPKPKDMYSRILIERLLKLAVDNHLIMLFFWMLFHSCLANVAELMOGDRE 360  
DB 301 VPTIONSKPKPKDMYSRILIERLLKLAVDNHLIMLFFWMLFHSCLANVAELMOGDRE 360  
QY 361 YRDWNNSSSVTYFMQNNNIPVHKWCI RHPYKPM LBRGSSKMAARTGVFLASAFPEHYLS 420  
DB 361 YRDWNNSSSVTYFMQNNNIPVHKWCI RHPYKPM LBRGSSKMAARTGVFLASAFPEHYLS 420  
QY 421 VPLNFRILMAFTGMAAQLPLANFVRPFGNGYNAAWLSLITIGPIAVLMTVHYVYL 480  
DB 421 VPLNFRILMAFTGMAAQLPLANFVRPFGNGYNAAWLSLITIGPIAVLMTVHYVYL 480  
QY 481 YEAPAAEA 488  
DB 481 YEAPAAEA 488

RESULT 9

AA54139

ID AA54139 strand; protein; 500 AA.

XX AA54139;

XX 27-MAR-2000 (first entry)

XX Acyl-CoA:cholesterol acyltransferase (ACAT)-like protein.

XX Acyl-CoA:cholesterol acyltransferase, ACAT, ACAT-like protein, sterol;

XX ester; triacylglycerol; fatty acyl-CoA; lipid composition; plant cell;

XX diacylglycerol acyltransferase, DAGAT, triglyceride; cancer; diabetes;

XX cardiopulmonary disease; heart failure; atherosclerosis; adipocytosis;

XX leukaemia; skin carcinoma; fibroblastoma; metabolic disorder; obesity;

XX abnormal lipid metabolism; abnormal fat absorption;

XX lipoprotein secretion; adipogenesis.

XX Rattus sp.

XX MO9963096-A2.

XX 09-DEC-1999.

XX 04-JUN-1999; 99WO-US012541.

XX 05-JUN-1998; 98US-0088143P.

XX 12-NOV-1998; 98US-0108389P.

XX (CALJ ) CALGENE LLC.

XX Lasener MW, Ruzinskiy DM;

XX WPI; 2000-105701/09.

XX N-PSDB; AA245385.

XX Novel polynucleotides used for modifying plant oil composition and for

XX developing products for treating e.g. cancer, diabetes, cardiopulmonary

XX disease or metabolic disorders.

XX Example 2; Fig 11; 89pp; English.

XX The present sequence represents an acyl-CoA:cholesterol acyltransferase

XX (ACAT) related protein. The ACAT-like protein is active in the formation

XX of a sterol, ester and/or triacylglycerol from a fatty acyl-CoA and

XX the lipid composition of plant cells. The ACAT-like protein has

XX diacylglycerol acyltransferase (DAGAT) activity, and so the synthesis of

XX triglycerides can be suppressed or increased using the DNA. The protein

XX can be used to produce plant oils with a modified triglyceride content.

XX The products can also be used to identify antagonists and agonists of

XX DAGAT activity. Such agonists and antagonists are particularly useful in

XX treating or ameliorating diseases associated with DAGAT activity.

XX including diseases associated with altered cellular diacylglycerol

XX concentration or PKC activity, including cancer, diabetes,

XX cardiopulmonary diseases e.g. heart failure, atherosclerosis,

XX adipocytosis, leukaemia, skin carcinoma, fibroblastoma, metabolic

XX disorders, obesity, diseases associated with abnormal lipid metabolism,

XX and diseases associated with abnormal fat absorption, lipoprotein

XX secretion and adipogenesis

XX Sequence 500 AA;

Query Match 88.6%; Score 2299.5; DB 3; Length 500;

Best Local Similarity 85.9%; Pred. No. 3.4e-241;

Matches 427; Conservative 26; Mismatches 31; Indels 13; Gaps 3;

QY 1 MGDGSSRRRTGSPSSSHGCGPAAAEBEVDAAAGPVGAAGAPAPAPKDGAGVG 51  
DB 1 MGDGSSRRRTGSPSSSHGCGPAAAEBEVDAAAGPVGAAGAPAPAPKDGAGVG 60  
QY 52 --NKDGDAGVSGHWELRCHRLQDSLFFSSDGSFNSYRGILNMCVMLILSNARLFLENLIK 109  
DB 52 --NKDGDAGVSGHWELRCHRLQDSLFFSSDGSFNSYRGILNMCVMLILSNARLFLENLIK 109



```

XX 31-OCT-2001; 2001WO-NZ000245.
PF
XX
XX 31-OCT-2000; 2000NZ-00507888.
PR
XX 06-DEC-2000; 2000NZ-00508662.
PR
XX
XX (GBOR/) GEORGES M A J.
PA (COPP/) COPPIETERS W H R.
PA (GRIS/) GRISART B M J.
PA (SNEI/) SNEEL R G.
PA (REID/) REID S J.
PA (FORD/) FORD C A.
PA (SPEL/) SPELMAN R J.
XX
XX Georges MJ, Coppieters WHR, Grisart BMJ, Snell RG, Reid SJ;
PI Ford CA, Spelman RJ;
XX
XX WPI; 2002-500128/53.
DR N-PSDB; AAD40350.
XX
XX
XX Determining genetic merit of a bovine with respect to milk composition
PT and volume for improved milk production, comprises determining the
PT diacylglycerol acyltransferase gene genotypic state of the bovine.
XX
XX
XX Claim 36; Page 68-69; 128pp; English.
XX
XX The invention relates to a method of genotyping bovine for improved milk
CC production traits which comprises determining the diacylglycerol
CC acyltransferase (DGAT1) genotypic state of the bovine, wherein the DGAT1
CC gene and polymorphisms have been found to be associated with such
CC improved milk production traits. The method is useful for selecting a
CC bovine having a desired DGAT1 genotypic state. It is also useful for the
CC identification and selection of a bovine having one of the polymorphisms
CC in its DGAT1 gene. Milk produced from selected bovine which is useful for
CC making a dairy product provides a beneficial health effect. An antibody
CC to the protein having DGAT1 activity is useful for inhibiting the
CC activity of bovine DGAT1 in a lactating bovine so as to modulate milk
CC production and/or milk solids content. DGAT1 nucleic acid and its
CC fragments are useful in the farming industry. They are also useful to
CC generate transgenic animals which are useful to investigate the molecular
CC basis of DGAT1 action and to test a substance for the ability to prevent,
CC slow or enhance DGAT1 activity. The present sequence is bovine DGAT1
CC protein. DGAT1 gene is located on chromosome 14
XX
XX
XX Sequence 489 AA;
SQ
Query Match 88.4%; Score 2292.5; DB 5; Length 489;
Best Local Similarity 88.0%; Pred. No. 1.9e-240;
Matches 434; Conservative 15; Mismatches 31; Indels 13; Gaps 3;
QY 1 MGDR---GSSRRRRRGRSSSHGGGPPAAEEVDDAAGPVGAAAGAPAPAPKDDDA 57
DB 1 MGDRCGAGGSGRRRRGRSPSIQGGSGPPAAEEV-----DVGAGGAPVADTDKDDV 54
QY 58 GVGSGHMEIRGRLQDLSLFSSSDSGSNYRGILNMCVWMLISNARLFENLIKSYILVDP 117
DB 55 DVGSGHMLRCHRLQDLSLFSSSDSGSNYRGILNMCVWMLISNARLFENLIKSYILVDP 114
QY 118 IQVNSLFLKDPHSPAPCLVIAANFAVAAPQVEKRLAVGALTQDAGLLHVAANIATTIC 177
DB 115 IQVNSLFLKDPHSPAPCLVIAANFAVAAPQVEKRLAVGALTQDAGLLHVAANIATTIC 174
QY 178 PPAAVVLVESITPPGSLIAAMAHITFLKLESYVDUNSWCR-----ARAAAGSGKAS 233
DB 175 PPAAVVLVESITPPGSLIAAMAHITFLKLESYVDUNSWCR-----ARAAAGSGKAS 234
QY 234 SAAPHTVSYDNLTYRDLVYFLFAPTLCEYELNFRSRRIRKRFILRLILEMFTQLOV 293
DB 235 GGAAGRTVSYDNLTYRDLVYFLFAPTLCEYELNFRSRRIRKRFILRLILEMFTQLOV 294
QY 294 GLIQGMVPTIQNSMKPFKMDYSHIIRLLKLAVPNHLIWLIFPYWLPHSCLNAVAELM 353
DB 295 GLIQGMVPTIQNSMKPFKMDYSHIIRLLKLAVPNHLIWLIFPYWLPHSCLNAVAELM 354

```

```

QY 354 QGDREFFRYDMWNSSEVTYFMQNNWNI PVHKNCIRHFFXKPMILRGSSKMMARTGTFLASAF 413
DB 355 QGDREFFRYDMWNSSEVTYFMQNNWNI PVHKNCIRHFFXKPMILRGSSKMMARTGTFLASAF 414
QY 414 FHEYLVSVPILRMFLMAFTGMAAQIPLAMFVGRFFQSGYNAVAWMLSLIIQPIAVLMVY 473
DB 415 FHEYLVSIPILRMFLMAFTGMAAQIPLAMVIGRFRGNYGNAVAWMLSLIIQPIAVLMVY 474
QY 474 HDYVLTANTEAPAA 486
DB 475 HDYVLTANTEAPAA 487
RESULT 12
AAE24979
ID AAE24979 standard; protein; 489 AA.
XX
XX AAE24979;
AC
XX
XX 22-OCT-2002 (first entry)
DT
XX
XX Bovine diacylglycerol acyltransferase (DGAT1) mutant, K232A.
DB
XX
XX Bovine; diacylglycerol acyltransferase; genotyping; milk production;
KW DGAT1; polymorphism; farming industry; transgenic; enzyme; mutant;
KW mutain.
XX
XX Bos taurus.
OS
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FT Misc-difference 232 /note= "Wild-type Lys substituted with Ala"
FT
XX
XX WO200236824-A1.
XX
XX 10-MAY-2002.
XX
XX 31-OCT-2001; 2001WO-NZ000245.
XX
XX 31-OCT-2000; 2000NZ-00507888.
PR 06-DEC-2000; 2000NZ-00508662.
XX
XX (GBOR/) GEORGES M A J.
PA (COPP/) COPPIETERS W H R.
PA (GRIS/) GRISART B M J.
PA (SNEI/) SNEEL R G.
PA (REID/) REID S J.
PA (FORD/) FORD C A.
PA (SPEL/) SPELMAN R J.
XX
XX Georges MJ, Coppieters WHR, Grisart BMJ, Snell RG, Reid SJ;
PI Ford CA, Spelman RJ;
XX
XX WPI; 2002-500128/53.
XX
XX Determining genetic merit of a bovine with respect to milk composition
PT and volume for improved milk production, comprises determining the
PT diacylglycerol acyltransferase gene genotypic state of the bovine.
XX
XX
XX Disclosure; Page; 128pp; English.
XX
XX The invention relates to a method of genotyping bovine for improved milk
CC production traits which comprises determining the diacylglycerol
CC acyltransferase (DGAT1) genotypic state of the bovine, wherein the DGAT1
CC gene and polymorphisms have been found to be associated with such
CC improved milk production traits. The method is useful for selecting a
CC bovine having a desired DGAT1 genotypic state. It is also useful for the
CC identification and selection of a bovine having one of the polymorphisms
CC in its DGAT1 gene. Milk produced from selected bovine which is useful for
CC making a dairy product provides a beneficial health effect. An antibody
CC to the protein having DGAT1 activity is useful for inhibiting the

```

CC activity of bovine DGAT1 in a lactating bovine so as to modulate milk  
 CC production and/or milk solids content. DGAT1 nucleic acid and its  
 CC fragments are useful in the farming industry. They are also useful to  
 CC generate transgenic animals which are useful to investigate the molecular  
 CC basis of DGAT1 action and to test a substance for the ability to prevent,  
 CC slow or enhance DGAT1 activity. The present sequence is bovine DGAT1  
 CC protein mutant used to illustrate the method of the invention. Note: This  
 CC sequence is not shown in the specification, however it is constructed  
 CC based on the sequence SEQ.ID.NO:2 shown in the sequence listing  
 CC (AAE24973)

XX Sequence 489 AA;

Query Match 88.1%; Score 2286.5; DB 5; Length 489;  
 Best Local Similarity 87.8%; Pred. No. 8.6e-240;  
 Matches 433; Conservative 15; Mismatches 32; Indels 13; Gaps 3;

QY 1 MGDR---GSSRRRTGSRPSSHGGGPPAAABEEVRDAAAGPDVGAAGAPAPAPAKDGA 57  
 DB 1 MGDRGAGGSRRRRTGSRPSIGGSGPAAABEEVR-----DVAGAGDAPVDTDXGDV 54  
 QY 58 GVGSGHMLRCHRLDLSFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP 117  
 DB 55 DVGSGHMLRCHRLDLSFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP 114  
 QY 118 IQVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTQAGLLHVALNATILC 177  
 DB 115 IQVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTQAGLLHVALNATILC 174  
 QY 178 FPAAVLVLESTITPVGSLALMAHTILFLKLSFYRDVNSWCR---ARAKASGKKAS 233  
 DB 175 FPAAVLVLESTITPVGSLALMAHTILFLKLSFYRDVNSWCRERRRAGAKAKALAGKKA 234  
 QY 234 SAAAPHVSYPDNLTLYRDLVYFLFAPTLCYELNPPRSRIRKRLRLLEMLFTQIQV 293  
 DB 235 GGAAGRTVSYPDNLTLYRDLVYFLFAPTLCYELNPPRSRIRKRLRLLEMLFTQIQV 294  
 QY 294 GLIQOMVPTIQNSMKPKMDYSRIIRLKLAVPNHLIWLIFPYWLFHSCINVAEIM 353  
 DB 295 GLIQOMVPTIQNSMKPKMDYSRIIRLKLAVPNHLIWLIFPYWLFHSCINVAEIM 354  
 QY 354 QFQDEPFYRDWNSSESVTFWQNNANI PVHKWCIRHFYKPMRLRGSSKMAARTGVLASAF 413  
 DB 355 QFQDEPFYRDWNSSESVTFWQNNANI PVHKWCIRHFYKPMRLRGSSKMAARTGVLASAF 414  
 QY 414 FHBTVSVPLMERLMAFTGMAAOIPLAMFVGRFQGNVNAVWLSLIGQPIVLMYV 473  
 DB 415 FHBTVSVPLMERLMAFTGMAAOIPLAMFVGRFQGNVNAVWLSLIGQPIVLMYV 474  
 QY 474 HDYVLANEAPPA 486  
 DB 475 HDYVLANEAPPA 487

RESULT 13  
 ABP96046  
 ID ABP96046 standard; protein; 489 AA.  
 AC ABP96046;  
 DT 07-MAY-2003 (first entry)  
 DE Bovine DGAT protein SEQ ID NO:4.  
 XX  
 XX ACyl CoA:diacylglycerol transferase; DGAT; enzyme; chromosome 14; bovine;  
 KM milk; meat marbling; low fat; polymorphic; SNP;  
 KM single nucleotide polymorphism.  
 XX  
 XX Bos taurus.  
 XX  
 XX WO2003004630-A2.  
 XX  
 XX 16-JAN-2003.  
 PD

XX  
 PF 05-JUL-2002; 2002WO-EP007520.  
 XX  
 PR 06-JUL-2001; 2001EP-00116412.  
 PR 13-MAY-2002; 2002US-0379412P.  
 XX  
 PA (ARBE-) ARBEITSGEMEINSCHAFT DEUT RINDERZUECHTER.  
 XX  
 PI Fries H, Winter A;  
 XX  
 DR WPI; 2003-239205/23.  
 DR N-PSDB; ABZ77045.  
 XX  
 PT New nucleic acid molecule comprising a sequence of an allele of a  
 PT polymorphic bovine acyl CoA:diacylglycerol transferase gene useful for  
 PT testing a mammal for its predisposition for fat content of milk and for  
 PT meat marbling.

Claim 13; Page 90-91; 91pp; English.

The present invention describes a nucleic acid molecule (NA) (i) encoding a bovine acyl CoA:diacylglycerol transferase (DGAT) contributing to or indicative for low fat content of milk and to low meat marbling (intramuscular fat content). Human DGAT is located to chromosome 8, and bovine DGAT is located to chromosome 14. (i) is useful for testing a mammal for its predisposition for fat content of milk and/or its predisposition for meat marbling. The method comprises analysing the gene encoding DGAT for nucleotide polymorphisms (e.g. single nucleotide polymorphisms (SNPs)) which are connected with the predisposition. The nucleotide polymorphisms are located in the coding region of the DGAT gene and result in substitution, deletion and/or addition of an amino acid sequence of the polypeptide which is encoded by the gene. The nucleic acid molecule has at the position 10433 and 10434 of the DGAT gene a guanine and a cytosine residue, at position 3343 a cytosine or guanine, 11030 a guanine, 11048 a cytosine or thymine and 11093 a thymine, which correlate with a predisposition for low fat content of milk and low meat marbling. The nucleic acid molecule has at the position 10433 and 10434 of the DGAT gene two adenine residues which correlate with a predisposition for high content of milk and high meat marbling. The nucleotide polymorphisms are located in a region which is responsible for the regulation of the expression of the product of the gene encoding DGAT. ABZ76924 to ABZ77045 and ABP96035 to ABP96046 represent sequences used in the exemplification of the present invention

XX Sequence 489 AA;

Query Match 87.8%; Score 2278.5; DB 6; Length 489;  
 Best Local Similarity 87.8%; Pred. No. 6.4e-239;  
 Matches 433; Conservative 14; Mismatches 33; Indels 13; Gaps 3;

QY 1 MGDR---GSSRRRTGSRPSSHGGGPPAAABEEVRDAAAGPDVGAAGAPAPAPAKDGA 57  
 DB 1 MGDRGAGGSRRRRTGSRPSIGGSGPAAABEEVR-----DVAGAGDAPVDTDXGDV 54  
 QY 58 GVGSGHMLRCHRLDLSFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP 117  
 DB 55 DVGSGHMLRCHRLDLSFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP 114  
 QY 118 IQVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTQAGLLHVALNATILC 177  
 DB 115 IQVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTQAGLLHVALNATILC 174  
 QY 178 FPAAVLVLESTITPVGSLALMAHTILFLKLSFYRDVNSWCR---ARAKASGKKAS 233  
 DB 175 FPAAVLVLESTITPVGSLALMAHTILFLKLSFYRDVNSWCRERRRAGAKAKALAGKKA 234  
 QY 234 SAAAPHVSYPDNLTLYRDLVYFLFAPTLCYELNPPRSRIRKRLRLLEMLFTQIQV 293  
 DB 235 GGAAGRTVSYPDNLTLYRDLVYFLFAPTLCYELNPPRSRIRKRLRLLEMLFTQIQV 294  
 QY 294 GLIQOMVPTIQNSMKPKMDYSRIIRLKLAVPNHLIWLIFPYWLFHSCINVAEIM 353

Db 295 GLIQMNVPAIQNSMKPFKMDYSRIVERLTKLAVPNHLIWLIFPYWLFHSCINAAVEIEM 354  
 QY QGDEEFYRDMMNSRVTYFMQNNMI PVHKICIRHFYKPMILRRGSSKMAARTGVLASAF 413  
 Db 354 QGDEEFYRDMMNSRVTYFMQNNMI PVHKICIRHFYKPMILRRGSSKMAARTGVLASAF 414  
 QY 414 FHELYSVPLRMFLMAFTGMAAQIPLAMFVGRFQNGYNAAVMLSLIIGOPVAVLMV 473  
 Db 415 FHELYSVPLRMFLMAFTGMAAQIPLAMVIGRFRNGYNAAVMLSLIIGOPVAVLMV 474  
 QY 474 HDYVYLVNTEAPAA 486  
 Db 475 HDYVYLVNTEAPAA 487  
 RESULT 14  
 ADCl6606  
 ID ADCl6606 standard; protein; 498 AA.  
 AC ADCl6606;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Mouse diacylglyceride acyltransferase (DGAT) protein sequence.  
 XX  
 KM diacylglycerol O-acyltransferase; diacylglyceride acyltransferase; DGAT;  
 KM triacylglycerol synthesis; intestinal fat absorption;  
 KM lipoprotein assembly; fat storage in adipocytes; milk production;  
 KM egg production; sperm maturation; hyperlipidaemia;  
 KM cardiovascular disease; obesity; diabetes; cancer; neurological disorder;  
 KM immunological disorder; sperm count; transgenic plant;  
 KM industrial feedstock; transgenic animal; anti-obesity; cardiant;  
 KM cyrostatic; antilipemic compound; mouse; murine; enzyme; EC 2.3.1.20.  
 XX  
 OS Mus musculus.  
 XX  
 PN MO967268-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 23-JUN-1999; 99WO-US014245.  
 XX  
 PR 24-JUN-1998; 98US-00103754.  
 PR 09-NOV-1998; 98US-0107771P.  
 XX  
 PA (RBC ) UNITV CALIFORNIA.  
 XX  
 PI Farese RV, Cases S, Smith S, Erickson SK, Sande E;  
 XX  
 DR WPI: 2000-147196/13.  
 DR N-PSDB; ADCl6605.  
 XX  
 PS Example 1; SEQ ID NO 5; 63pp; English.  
 XX  
 CC This invention relates to novel diacylglycerol O-acyltransferase  
 CC (diacylglyceride acyltransferase; DGAT; EC 2.3.1.20) nucleotide and protein  
 CC sequences. The DGAT enzyme is involved in triacylglycerol synthesis, and  
 CC is important in higher eukaryotes for intestinal fat absorption,  
 CC lipoprotein assembly, fat storage in adipocytes, milk production, and  
 CC possibly egg production and sperm maturation. The DGAT nucleotide  
 CC sequences may be useful for identifying DGAT homologues, as a source of  
 CC probes and primers, for the identification of expression patterns in  
 CC biological specimens, for the preparation of cell or animal models of  
 CC DGAT function, and for the preparation of in vitro models of DGAT  
 CC function. The invention may also be used for diagnostic screening, for  
 CC the treatment of pathological conditions, and for treatment of/or created  
 CC related to DGAT expression or activity. Diseases diagnosed and/or treated  
 CC include hyperlipidaemia, cardiovascular disease, obesity, diabetes,  
 CC cancer, neurological disorders and immunological disorders. DGAT may also  
 CC be administered to males to increase sperm count. Also, transgenic plant

CC with altered levels of DGAT can be used to produce novel oils for  
 CC consumption by humans, or for use as industrial feedstocks. Transgenic  
 CC animals with altered levels of DGAT can be used as sources for a variety  
 CC of different food and industrial products in which the triglyceride  
 CC content is specifically tailored. The invention may enable development of  
 CC anti-obesity, cardiant, cyrostatic or antilipemic compounds. The present  
 CC sequence is that of the mouse DGAT protein of the invention. Note: The  
 CC present sequence does not appear in the specification but was obtained by  
 CC the index from GenBank.  
 XX  
 SQ Sequence 498 AA;  
 Query Match 87.8%; Score 2278.5; DB 3; Length 498;  
 Best Local Similarity 85.1%; Pred. No. 6.6e-239;  
 Matches 421; Conservative 29; Mismatches 34; Indels 11; Gaps 3;  
 QY 1 MGDR---GSSRRRTGSRPSSHGSGGPPAAAEVYDAAAGDVGAGAPAPAP----- 51  
 Db 1 MGDREGAGSSRRRRRGSRVSVQGSQSGPYEEDVYDAVSDPLGAGGAPAPAPAPATRR 60  
 QY 52 NKDGDAGVSGHWEIQRRLDLSLSSSDSGFSNRYGLNMCVWMLISNARLPLENLKY 111  
 Db 61 DKDERTVGDGYMDRCHRLDLSFSSDSGFSNRYGLNMCVWMLISNARLPLENLKY 120  
 QY 112 GILVDPPIQVSLFLKDPHSPAPCLVIAANFAVAAPQVBRRLAVGALTQAGLLHAVN 171  
 Db 121 GILVDPPIQVSLFLKDPHSPAPCLVIAANFAVAAPQVBRRLAVGALTQAGLLHAVN 180  
 QY 172 LATIICPPAAVVLVESITPVGSLIAHMAHTILFKLFPSYDVNSWC--RRAPAKAAG 229  
 Db 181 LATIICPPAAVVALVESITPVGSVFALASVIMFKLTSYDVNLMCRQRRKAKAVSTG 240  
 QY 220 KKASAAAPHVSVPDNLTVDLYVFLFAPLTCVLELNPSPRIRKRLRLRIEMLFPT 289  
 Db 241 KKVSGAAQAQVSPDNLTYRDLTYFLFAPLTCVLELNPSPRIRKRLRLRLRIEMLFPT 300  
 QY 290 QLOVGLIQMNVPTIQNSMKPFKMDYSRIIERLTKLAVPNHLIWLIFPYWLFHSCINAV 349  
 Db 301 QLOVGLIQMNVPTIQNSMKPFKMDYSRIIERLTKLAVPNHLIWLIFPYWLFHSCINAV 360  
 QY 350 AELMQFGDREFFYRDMMNSRVTYFMQNNMI PVHKICIRHFYKPMILRRGSSKMAARTGVPL 409  
 Db 361 AELLQFGDREFFYRDMMNSRVTYFMQNNMI PVHKICIRHFYKPMILRRGSSKMAARTGVPL 420  
 QY 410 ASAFHELYSVPLRMFLMAFTGMAAQIPLAMFVGRFQNGYNAAVMLSLIIGOPVAV 469  
 Db 421 TSAPFHELYSVPLRMFLMAFTGMAAQIPLAMVIGRFRNGYNAAVMLSLIIGOPVAV 480  
 QY 470 LMYVHDYVYLVNTEAP 484  
 Db 481 LMYVHDYVYLVNTEAP 495  
 RESULT 15  
 ADS00390  
 ID ADS00390 standard; protein; 498 AA.  
 AC ADS00390;  
 XX  
 DT 16-DEC-2004 (first entry)  
 XX  
 DE Mouse diacylglycerol acyltransferase 1.  
 XX  
 KM Antiartherosclerotic; Antilipemic; Antidiabetic; Anorectic; Cardiant;  
 KM Diacylglycerol acyltransferase-Inhibitor-1;  
 KM diacylglycerol acyltransferase 1; glucose; cholesterol; triglyceride;  
 KM lipid metabolism; cholesterol metabolism; atherosclerosis;  
 KM hyperlipidaemia; diabetes; type 2 diabetes; obesity;  
 KM cardiovascular disease; mouse; enzyme.  
 OS Mus musculus.  
 XX  
 PN US2004185559-A1.





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

## OM protein - protein search, using bw model

Run on: May 5, 2006, 12:29:50 ; Search time 41 Seconds  
(without alignments)  
1145.214 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MCDRGSSRRRTGSRPSSHG.....VLMYVDYVLYNLEAPAAEA 488

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1106	42.6	496	2	probable sterol O-
2	791.5	30.5	520	2	diacylglycerol O-a
3	661	25.5	441	2	diacylglycerol O-a
4	467	18.0	231	2	sterol O-acyltrans
5	371	14.3	540	1	sterol O-acyltrans
6	345	13.3	550	1	sterol O-acyltrans
7	344.5	13.3	642	2	probable membrane
8	341.5	13.2	437	2	probable sterol O-
9	300	11.6	537	2	probable sterol O-
10	297	11.4	467	2	hypothetical prote
11	281.5	10.9	610	2	hypothetical prote
12	131.5	5.1	305	2	sterol O-acyltrans
13	121.5	4.7	560	2	probable membrane
14	117.5	4.5	452	2	sn-glycerol-3-phos
15	117.5	4.5	452	2	sn-glycerol-3-phos
16	116	4.4	609	2	probable membrane
17	113.5	4.4	452	1	glycerol-3-phospha
18	110	4.2	659	2	hypothetical prote
19	107.5	4.1	475	2	hypothetical prote
20	106.5	4.1	478	2	probable polysacch
21	106	4.1	495	2	probable alginat
22	106	4.1	527	2	alginat O-acetyla
23	100.5	3.9	478	2	alginat O-acetyla
24	99.5	3.8	567	2	menaquone biosyn
25	98.5	3.8	247	2	hypothetical prote
26	97	3.7	473	2	closely related to
27	96	3.7	267	1	corticotropin / 11
28	95.5	3.7	452	2	glycerol-3-phospha
29	95.5	3.7	519	2	tachykinin recepto

30	95.5	3.7	807	2	oxyterol-binding
31	95	3.7	426	2	Bcl protein precu
32	94.5	3.6	238	2	hydrogenase (EC 1.
33	94.5	3.6	457	2	NADH2 dehydrogen
34	94.5	3.6	583	2	hypothetical prote
35	93	3.6	416	2	hypothetical 45.0
36	93	3.6	572	2	alpha-1A-adrenerg
37	93	3.6	932	2	suvar(3)7 protein
38	92.5	3.6	287	2	hypothetical prote
39	91.5	3.5	403	2	Tetracycline resis
40	90.5	3.5	299	2	probable hydrogena
41	90.5	3.5	395	2	hypothetical prote
42	90.5	3.5	600	2	NADH2 dehydrogen
43	90.5	3.5	1911	2	calcium channel al
44	90.5	3.5	2410	2	cell wall alpha-gl
45	90	3.5	659	2	hypothetical prote

## ALIGNMENTS

RESULT 1  
T19027  
probable sterol O-acyltransferase (EC 2.3.1.26) H19N07.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #ext\_change 09-Jul-2004  
C/Accession: T19027, T23106  
R.Kerhaw, J.  
submitted to the EMBL Data Library, June 1996  
A/Reference number: Z19061  
A/Accession: T19027  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-496 <MT>  
A/Cross-references: UNIPROT:Q45245; UNIPARC:UPI000007F625; EMBL:Z75526; PIDD:CAA99773.1  
R.Dobson, R.  
submitted to the EMBL Data Library, March 1997  
A/Reference number: Z19678  
A/Accession: T23106  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-496 <MT>  
A/Cross-references: UNIPARC:UPI000007F625; EMBL:Z92835; PIDD:CAB07399.1; GSPDB:GN00023;  
A/Experimental source: clone H19N07  
C/Genetics:  
A/Gene: CBSP:H19N07.4  
A/Map position: 5  
A/Introns: 5/3; 43/2; 77/3; 139/1; 171/3; 456/1  
C/Superfamily: sterol O-acyltransferase  
C/Keywords: acyltransferase; coenzyme A

Query Match 42.6%; Score 1106; DB 2; Length 496;  
Best Local Similarity 43.7%; Pred. No. 2.2e-83;  
Matches 220; Conservative 89; Mismatches 155; Indels 40; Gaps 6;

QY 6 SRRRRRTGSRPSSHGCGPAAAEVRODAAGDPVGAAGDAAPAPNPDGAGVSGHME 65  
DB 5 TGRRRR---QPSRTSGSLASSRR-----SFPANQNSRRKSSKMRGCE 46  
QY 66 LRCHRLQDLSFSSDGFNSYRGILNMCVMILISNARLFLENLIRYGLIVDPQVVSJFL 125  
DB 47 KVMHTAQDSLSTSGMWNFRGFFRLSLTLVLSNGRVALENVIRKGLITPLQWISTFV 106  
QY 126 KDPH--SWPAPCVIAANVPAVAAFQVEKRLVAGALTBOGLLIHVNLAITLCPPAV 182  
DB 107 EHHYIWSGMPNIALICSNIOILSVGMEKILERGMNGFAAVFTSLVIAHLYIPVVV 166  
QY 103 VLVVSIIPVCSILALMAHTILFLKLFYRDVNSCRBARAPAAAGKASSAA----- 236  
DB 167 TLTHKMKPLMSVVMGIVIALKFLFISYGHNVWARRRKITTLKQTQVTDLAKKTCDP 226  
QY 237 -----APHTVS--YPDNLTFYRDLYYFLFAFTLCYEINFPSPRIKRFLRLILE 284



sterol O-acyltransferase homolog F27F23.26 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T01294  
 R:Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, May 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC F27F23 genomic sequence.  
 A:Reference number: 214177  
 A:Accession: T01294  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-231 <R0U>  
 A:Cross-references: UNIPROT:Q9SLD2; UNIPARC:UPI000017B01B; EMBL:AC003058; NID:g3135250;  
 C:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Map position: 2  
 A:Introns: 44/3; 70/3; 104/3; 150/2; 177/3; 184/3; 205/3  
 A:Note: F27F23.26

Query Match 18.0%; Score 467; DB 2; Length 231;  
 Best Local Similarity 44.0%; Pred. No. 4.6e-31;  
 Matches 103; Conservative 42; Mismatches 65; Indels 24; Gaps 8;

QY 193 GSLALMAHTILFLKLFYSYR---DVNSMCRBARAKAAGKASAAAPHTVSYPDNLT 248  
 DB 11 GVTLMMLT-CIVMLKLVYAHTSYDIRS-----LANAADKANP-----EVSY--YYS 54  
 QY 249 YRDLYFLFAPPTLCYELNPPSPRIKRLRLILEMFLFTQLQVGLIQOMVPTIQNSM 308  
 DB 55 LKSLAVFVWAPPLCQPSRSACIRKGVARQAFALVITFGMEPIIDIOYINPFRNKK 114  
 QY 309 KPEK-DMDYSRIERLKLKAVPNHILWLFYFWLFFHSCNAVAELMOFGDREPYRDWNS 367  
 DB 115 HPLKEDLYA--IERVLKLSVENLYVWLCMFYCFHLMNITLIELCFDREPYKDWMA 172  
 QY 368 ESVTYFMQWNIPVHKMCIRHFYKPMRLRRGSSKMMARTGVFLASAFHE-YLVS 420  
 DB 173 KSVGDYMRMMNPVHKMVRHITYPCLRGSKIPKTLAIIIAFLVSAVFHEVITLS 226

## RESULT 5

sterol O-acyltransferase (EC 2.3.1.26) - mouse  
 N:Alternate names: acyl-coenzyme A cholesterol acyltransferase  
 C:Species: Mus musculus (house mouse)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: J49454; J04617  
 R:Uelmen, P.J.; Oka, K.; Sullivan, M.C.; Chang, T.  
 J. Biol. Chem. 270, 26192-26201, 1995  
 A:Title: Molecular cloning of mouse ACAT.  
 A:Reference number: 149454; MUID:96064687; PMID:7592824  
 A:Accession: J49454  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-540 <RES>  
 A:Cross-references: UNIPROT:061263; UNIPARC:UPI000002816F; GB:I42293; NID:g1066809; PIDN  
 R:Green, S.; Steinberg, D.; Quehenberger, O.  
 Biochem. Biophys. Res. Commun. 218, 924-929, 1996  
 A:Title: Cloning and expression in Xenopus oocytes of a mouse homologue of the human acy  
 A:Reference number: J04617; MUID:96158986; PMID:8579615  
 A:Accession: J04617  
 A:Molecule type: mRNA  
 A:Residues: 1-194, 'R', 196-540 <GRE>  
 A:Cross-references: UNIPARC:PI0000170CD5; GB:S81092; NID:g1478335; PIDN:AAB36050.1; PID  
 A:Experimental source: peritoneal macrophages  
 C:Comment: This enzyme helps maintain cellular cholesterol homeostasis by catalyzing est  
 Y cholesterol and oxygenated sterole.  
 C:Genetics:  
 A:Gene: ACAT  
 A:Map position: 1  
 C:Superfamily: sterol O-acyltransferase  
 C:Keywords: acyltransferase; cholesterol; coenzyme A; endoplasmic reticulum; transmembr

Query Match 14.3%; Score 371; DB 1; Length 540;  
 Best Local Similarity 27.1%; Pred. No. 1e-22;  
 Matches 126; Conservative 70; Mismatches 167; Indels 102; Gaps 19;

QY 71 LDDSLFSSDSGSSNRGLIIMCVMLIL-----SNARLFE-NLIKGLIVDI 118  
 DB 118 LIDELFEVD---HRTTYHMFILILFLVSLTIVDYIDEGRLVLEFLLVAFGKFP 173  
 QY 119 QY-----VSLFKDPHSWPAPCLVIAAVFAVAFOVERKRLVGLTEQGL 166  
 DB 174 VITWVWAMFLSTLSIPFLFOP-W-----AHGYSKSHPLIYSLV-----HGL 216  
 QY 167 LHVANLATICPPAAVVLVESITPVGSLALMAHTILFLKLFYSYRDVNSMCRBARAKA 226  
 DB 217 FLVPLQGL-VLGFVPTVYVLAAYLTPASRLITLLEGIRLMK-----ASHFRENIPRL 269  
 QY 227 SAGKASAAADPTVSYPDNLTYYRLYFLFAPPTLCYELNPPSPRIKRLRLILEM 285  
 DB 270 NAAKSSKSDPLPTVN-----OYLVELFAPTLIYRDNYRPTVVRGVAMOFLOVF 321  
 QY 286 --LFTQLQVGLIQOMVPTIQNSMCRBARAKAAGKASAAAPHTVSYPDNLT 337  
 DB 322 GCLFVYVY--IFERLCAPLFRNIKQEP8-----ARVLVLCVNSILFGLILFUS 370  
 QY 338 FYWLFHSCINAAVAELMOFGDREPYRDWNSBSVTFMQWNIPVHKMCIRHFYKPMRLRG 397  
 DB 371 FFAFLHCHWLNAPAEMLRQDRMFYKDWNSISYSYRTNVAVDMLYYVYKOLLWF 430  
 QY 398 SSKM--MARTGVFLASAFHEYLVSPRL---MFLMAFTGMAQIPLANFVGRFQ 450  
 DB 431 SKRFSAAMLAVFALSAVHEVALAICLSFYFVFLFMPFGM-----AFNIVDSRRK 486  
 QY 451 NYGNAVWLSLIIGPIAVLMYVHDYV-----LNTAP 484  
 DB 487 PIVNIMWASLPLVGLITLCFYSQEMVARIQHCPPLNPTFLDYVP 531

## RESULT 6

sterol O-acyltransferase (EC 2.3.1.26) - human  
 N:Alternate names: ACAT; acyl-coenzyme A cholesterol acyltransferase  
 C:Species: Homo sapiens (man)  
 C:Date: 28-May-1999 #sequence\_revision 28-May-1999 #text\_change 09-Jul-2004  
 C:Accession: A59038; A48026  
 R:Chang, C.C.Y.; Chang, T.Y.  
 submitted to Genbank, May 1999  
 A:Description: Molecular cloning and functional expression of human acyl-coenzyme A:cho  
 A:Reference number: A59038  
 A:Contents: correction  
 A:Accession: A59038  
 A:Molecule type: mRNA  
 A:Residues: 1-550 <CH2>  
 A:Cross-references: UNIPROT:P35610; UNIPARC:UPI0000135B5D; GB:I21934; NID:g4878021; PIDN  
 R:Chang, C.C.Y.; Huh, H.Y.; Cadigan, K.M.; Chang, T.Y.  
 J. Biol. Chem. 268, 20747-20755, 1993  
 A:Title: Molecular cloning and functional expression of human acyl-coenzyme A:cholester  
 A:Reference number: A48026; MUID:94012607; PMID:8407899  
 A:Accession: A48026  
 A:Molecule type: mRNA  
 A:Residues: 1-206, 'R', 208-550 <CH1>  
 A:Cross-references: UNIPARC:UPI000014346C; GB:I21934  
 C:Genetics:  
 A:Gene: GDB:SOAT; STAT; ACAT  
 A:Cross-references: GDB:251696; OMIM:102642  
 A:Map position: 1q25-1q25  
 C:Function:  
 A:Description: catalyzes the esterification of cholesterol by acyl-CoA  
 A:Pathway: cholesterol metabolism  
 A:Note: helps maintain cellular cholesterol homeostasis; plays a role in the development  
 C:Superfamily: sterol O-acyltransferase  
 C:Keywords: acyltransferase; cholesterol metabolism; coenzyme A; endoplasmic reticulum;  
 F/409,491/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.3%; Score 345; DB 1; Length 550;  
Best Local Similarity 26.8%; Pred. No. 1.4e-20;  
Matches 121; Conservative 73; Mismatches 182; Indels 76; Gaps 19;

QY 71 LQDSLFSSDGSFNSYRGLNMCVNLISNARLPLENTIKYGLVDPQVVS-LFLKDP-128  
DB 128 LDELLEVD---HRTIYHMFALLILFISTLVVDYDEGRVLEFSLISYAGKEPPT-183  
QY 129 --HSPAPCLVIAAVFAVAAFOVEKRLAVGALTQO-----AGLLHVNANLNTIICF-178  
DB 184 VVMTW---WIMFSLTFSPVPF-LFOHMYATGYSKSHPLIRSLFHFGLFMITQIG-VLGF-237  
QY 179 PAAVLLVESITPVPSGLLALMAHTILFLKFSYRDVNSWCRPARAKAASAGKASSAAP-238  
DB 238 GPTVYVAVATLPPARFI-----IIFGIRFVMAKSHFRENVRVLSAEEKS-----287  
QY 239 HTVSYPDNLTYRDLVYFLPAPTLCYELNPPSPRIKRFLLRIEML--FTQLOVGLI-296  
DB 288 -TVPIP--TVQOYLYFLFAPTLIYRDSYPRNPVTWVGVAKFAQVGFCEFFVYVI--F-341  
QY 297 QQMWVPTIQN-SMKPFKMDYSRIIERLLKLAV-----PNHLIMLIFPYWLPHSCLNVA-350  
DB 342 ERLCALPRLNKKQEPFS-----ARVAVLCVENSILPGLILFLTFPAFLHCWLNAP-393  
QY 351 ELMQFDEREFYRDWNNSESVTYFQNNWNIPIVHKWCIHFYKPEMLRRGSSKM--MARTGVF-408  
DB 394 EMLRGDMFYKDWNNSSYSYRYTNVNVVDMLYYAKOFLWFFSRRFSSAAMLVF-453  
QY 409 LASAFHEYLVSVPJR---MFLMAFTGMAQIPLAMFVGRFQNGYGNAAVWLSLI-463  
DB 454 AVSAVAVHEYALAVCLSFYFVFLVFMFEGM---AFNFIYVDSRKKPIWVLTMTSFL-509  
QY 464 GQPIAVLMVYHDYV-----LNYEAR-484  
DB 510 GNGVLLCFYSQEWYARRHCPLKNPTFLDYRP-541

RESULT 7  
S63350  
probable membrane protein YNR019w - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein N3206  
C/Species: Saccharomyces cerevisiae  
C/Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 09-Jul-2004  
C/Accession: S63350  
R/Prohl, T.M.  
submitted to the Protein Sequence Database, April 1996  
A/Reference number: S63346  
A/Accession: S63350  
A/Molecule type: DNA  
A/Residues: 1-642 <POH>  
A/Cross-references: UNIPROT:P53629; UNIPARC:UPI0000125DBD; EMBL:Z71634; NID:G1302503; PI  
A/Experimental source: strain S288C  
A/Genetic:  
A/Gene: SGD:ARE2  
A/Cross-references: SGD:S0005302; MIPS:YNR019w  
A/Map position: 14R  
C/Superfamily: probable membrane protein YCR048w  
C/Keywords: transmembrane protein  
F/213-235/Domain: transmembrane #status predicted <TM1>  
F/293-311/Domain: transmembrane #status predicted <TM2>  
F/321-337/Domain: transmembrane #status predicted <TM3>  
F/435-451/Domain: transmembrane #status predicted <TM4>  
F/483-499/Domain: transmembrane #status predicted <TM5>  
F/622-638/Domain: transmembrane #status predicted <TM6>

Query Match 13.3%; Score 344.5; DB 2; Length 642;  
Best Local Similarity 28.6%; Pred. No. 1.9e-20;  
Matches 98; Conservative 66; Mismatches 138; Indels 41; Gaps 13;

QY 159 LTEQGLLHVNANLTLCPFAAVLVV--SITVGSLLAMAHITLFLK--LFSYD-213  
DB 311 LLEN---ILKHLWLSKIFLFLHSLVLLMKMHSFAFYNGYLMGIRKELOPSKSLAKLYKDS-367

QY 214 VNSMCRBARAKAASAGKK-----ASSAAAPHTVSPDNLTYRDLVYFLFAPTLCE-264  
DB 368 IND-----PVIGALBKSCFCSPSELSSQSLSDQTOGFNNISAKSPFMFTMTFLIYO-421  
QY 265 LNFPSPIRKRFLLRLLEN---LFTQLOVGLIQQMWVPTIQNSMKPFKMDYSRIE-321  
DB 422 IEYPTKEIRFSVYLEKICALFGTIFLMMIDAQIL---MYVANRAL-AVANSSEWTGLD-477  
QY 322 RLKLKLA-----VPHNL-LIIFPYWPHSCLNVAELMQFDEREFYRDWNNSESVTYF-374  
DB 478 RLKLKGLLVIVGFIYMTYLDLIDALINCYAELTRFGDXYFYGDWNNCVSMADFS-537  
QY 375 QNNWIPVHKWCIHFY-KPMLRRGSSKMMARTGVFLASAFHEYLVSVPRLMAFTG-433  
DB 538 RIMNIPVHKFLLRHVYHSSMSFPLNKSQATLMTFFLSSVHLEAMVYIFKRLRYLFF-597  
QY 434 MMAQIPLAMFVGRFQGN---YGNAAVWLSLIQPIAVLMVY-473  
DB 598 QMLQMPVALVNTIKFMRNRTIIGVIFWLIGICMGSPVCTLYL-640

RESULT 8  
T41684  
probable sterol o-acetyltransferase 2 - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T41684  
R/Wedder, H.; Dueserhoef, A.; McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, September 1999  
A/Reference number: Z21742  
A/Accession: T41684  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-472 <MED>  
A/Cross-references: UNIPROT:Q9U082; UNIPARC:UPI0000069EC3; EMBL:AL117183; PIDN:CA854864.  
A/Experimental source: strain 972h-; clone pl pLE11  
A/Genetic:  
A/Gene: SPDB:SPCP1E11.05C  
C/Superfamily: probable membrane protein YCR048w

Query Match 13.2%; Score 341.5; DB 2; Length 472;  
Best Local Similarity 26.7%; Pred. No. 2.3e-20;  
Matches 115; Conservative 73; Mismatches 173; Indels 69; Gaps 15;

QY 88 ILNMCVNLISNARLPLEN-----LIKY-GILVD---PIQVSLFLKDPHS-130  
DB 66 VLFVVAASIMTFMS-FLNFEELGRPVYGTIFKTFQSNLDDLAKADLAMSGMFL-----118  
QY 131 WPAPCLVIAAVFAVAAFOVEKRLAVGALTQOAGLLHVNANLTLCPFAAVL--LVES-188  
DB 119 -----LAPFQKIFALGYL-RWYGLGVLYS1-LILFLSHCVLRCLSN-161  
QY 189 ITPVGSLLAMAHITLFLKFSYRDVNSW---CRRARAKAASA-----GKKASSAAPH-239  
DB 162 WSWTRHAFILHSWILMKLHSYNNVGMWYSYCVHSLKQSKTDLDDDRSSVEFYE-221  
QY 240 TV-----SYPNLTGYRDLVYFLPAPTLCYELNPPSPRIKRFLLRIEMLFTQLOVG-294  
DB 222 CINHGNTYFENLITPNLDFLWPSLCYQLYYRTAHVRVHYLIECLGTFGCFLVY-281  
QY 295 LIQQMWVPTIQNSMKPFKMD--DVS-----RIERLLKLAVPNHLIWLIFPYWPHSCL-346  
DB 282 ISDHMVVVLAKAIRTIIEAPEDASATYFARLGHVTFALNFPFLSFLVFWYFEGVC-341  
QY 347 NAAVLEMQFDEREFYRDWNNSESVTYFQNNWNIPIVHKWCIHFYKPEMLRRGSSKMMARTG-406  
DB 342 NFSSEITTFARNFYDDMMNCMTWDOFARTNKKRPVYFLKLVHVVYV-LNSFWSKSLSTFF-400  
QY 407 VFLASAFHEYLVSVPRLMAFTGMAQIPLAMFVGRFQGNV---GNAAVWLSLI-463  
DB 401 TFFVSSVTLHELVMGITLIKIGYGLFQMTQIPIYIIORQKRVRRHRLGNAFNFSTII-460



Matches 99; Conservative 70; Mismatches 171; Indels 61; Gaps 14;  
QY 83 SNYGI-----LNMCVMLLSNARLFLENLTKYGLVDPQVSLFLKDPHSMRPC 135  
DB 184 SNFGIYVFAWMFLGMI-----RCCTDYASYSAMNKLTVQWTTDLFTIAMLD 236  
QY 136 LVIANVAV--AAFOVEKRLAVGALTQAGLLHVMANLARI-LCPFAVALL-VESTTP 191  
DB 237 LAMFLCTFFVFMHVLVKRIINMKWT--GFAVASIFELAFIPVTFIVVYFEDNMWTR 294  
QY 192 VGSLLAMAHITLF-LKLFSTRDVNSWCRARAKAASGK--KASSAAPHT----- 240  
DB 295 I-----FLFLHSYVFEMKSHSFAYNGYMDIKQBLEYSKQLOKESLSEPTREILQKS 350  
QY 241 -----VSVDNLTFRLLYFLFAPTLCELANPFRSRIRIKRFLRRLML 286  
DB 351 CDFCLFELNYOTKDNDFNNISCSNPFMECLFPLVLYQINPRISIRIMRYLEVVCAL 410  
QY 287 FETOLQVGLIOQWVPTIQNSMKPFKMDYSRIE-----RLKLAVPNH-LIMLIFPY 339  
DB 411 GTIFLAWYTAQFPMHVPVAMRCIQPHNTPTFGMIPATQEWPHLLPDMIPGTVLWMLFY 470  
QY 340 WLFHSCINAVALEMQGDEPFYRDWNSSEVTFEWNQNNIPVHKVCIRHFYKMLRGS- 398  
DB 471 MIMDALLNCVAVELTFPADRYFGYDWMNCVSFEFESRINNVVPHKFLRLHVVHSSW--GAL 528  
QY 399 --SKMARTGVFLASAFHEHYVSVPLMFRMAFTGMAQ 437  
DB 529 HLSQSQATLFTFFLSAVFHEMAMFAIFRRVGRYLELFWQSQ 569

RESULT 12  
147040  
sterol O-acyltransferase (EC 2.3.1.26) - rabbit (fragment)  
N:Alternate names: acyl-coenzyme A cholesterol acyltransferase  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 09-Jul-2004  
C:Accession: 147040  
R:Page, M.E.; Schlitz, P.A.; Rea, T.J.; Dematovic, R.B.; Kieft, K.; Bisgaler, C.L.; Newt  
J. Lipid Res. 36, 823-838, 1995  
A:Title: Tissue specific changes in acyl-CoA: cholesterol acyltransferase (ACAT) mRNA le  
A:Reference number: 147040; MUID:95341197; PMID:7616126  
A:Accession: 147040  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-505 <PAB>  
A:Cross-references: UNIPROT:Q95214; UNIPARC:UPI00008764A; GB:S78180; NID:q1515471; PIDN  
C:Comment: This enzyme helps maintain cellular cholesterol homeostasis by catalyzing est  
C:Superfamily: sterol O-acyltransferase  
C:Keywords: acyltransferase; cholesterol; coenzyme A; endoplasmic reticulum; transmembra

Query Match 5.1%; Score 131.5; DB 2; Length 305;  
Best Local Similarity 24.3%; Pred. No. 0.0029;  
Matches 63; Conservative 43; Mismatches 106; Indels 47; Gaps 12;  
QY 71 LODSIFSSDSGFSNRYGLINMCVWMLI-----SNARLFE-NLKYGLVDP 118  
DB 66 LIDELFEVD---HRTTYHMFALILFISTLVVDYIDBGRLLVLENLSYAGKGLPT 121  
QY 119 QVNSLFLKDPHSMRPCVIA--ANVFAVAFQVEKRLAVGALTQAGLLHVMANLARI 176  
DB 122 VVMTWMTMFLSTLSIPFLFGHWANGYSSKSHPLWYSLF-----HGLTFVWQLG-IL 173  
QY 177 CFPAAVLLVESITPVGSLALMAHTILFLKLFSTRDVNSWCRARAKAASGKASSAA 236  
DB 174 GFGPTVYLAATLPASRPVILEOIRLIMK-----ASHFVRENVPRYLANSAKSSS-- 225  
QY 237 APHTSYSDNLTFRLLYFLFAPTLCELANPFRSRIRIKRFLRRLML 286  
DB 226 --TVPIP--TVNOYLFLFLFAPTLFRDSYPRTPFRKGVYAMQAPQVFGCLFYVY-- 277  
QY 294 GLIOQWVPTIQN-SMKPF 311

DB 278 -IFERLCAPLFRNIKOEPP 295  
RESULT 13  
564091  
Probable membrane protein YGL084C - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein G3195  
C:Species: Saccharomyces cerevisiae  
C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004  
C:Accession: 564091  
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: 564071  
A:Accession: 564091  
A:Molecule type: DNA  
A:Residues: 1-560 <RIB>  
A:Cross-references: UNIPROT:P53154; UNIPARC:UPI000013B0BC; EMBL:272606; NID:q1322606; PI  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:GUP1; MIPS:YGL084C  
A:Cross-references: SGD:S0003052  
A:Map position: 7L  
C:Keywords: transmembrane protein  
F:47-63/Domain: transmembrane #status predicted <TM1>  
F:104-120/Domain: transmembrane #status predicted <TM2>  
F:132-148/Domain: transmembrane #status predicted <TM3>  
F:324-340/Domain: transmembrane #status predicted <TM4>  
F:356-372/Domain: transmembrane #status predicted <TM5>  
F:433-449/Domain: transmembrane #status predicted <TM6>  
F:453-469/Domain: transmembrane #status predicted <TM7>  
F:492-508/Domain: transmembrane #status predicted <TM8>  
F:529-545/Domain: transmembrane #status predicted <TM9>

Query Match 4.7%; Score 121.5; DB 2; Length 560;  
Best Local Similarity 19.0%; Pred. No. 0.04;  
Matches 101; Conservative 81; Mismatches 180; Indels 169; Gaps 28;  
QY 49 PAPNKGADGAVGSGSHW-----LRCRL--QD 73  
DB 22 PSPKKDASTTKRPSLMKTKERFYIAFLVVPVPLMFAGLQASSPENNVARYERLLSQ 81  
QY 74 SLFSS--DSGFSNRYGLINMCVWMLLSNARLFLENLTKYGLVDPQVSLFLKDPHSM 131  
DB 82 WLFGRKVDNSDSQRFEDNFDALLSVLMTVTSIKRILVSTNTTKRFDILF----- 134  
QY 132 PAPCLVIAANFANVAAPOVEK-----RLAVGALTQAGLLHVMANLARI-LCPFAVALL 185  
DB 135 -----GLFLVAAGVNSIRILAHMLIYAI--AHVKNRRRIATISIMYIGISTL 183  
QY 186 VES-----ITPVGSILALMA-----HTILFLKFSYR--DVNSWCRARA 223  
DB 184 FINDNFRAYPPGNCISFLSPDHWYRGILPRMDVFNFNTLRVSLYNLDFERENLQKK 243  
QY 224 KAAS-AGKASSA-----AAPHY-----SYDNLTYRLLYFLFAPTLCELANF 267  
DB 244 KSPSYSEKSAKSAIILNERARLTAAHPIDQYSIMNAYIYVYTP--FIAGPIITTFNDYV 301  
QY 268 PRS-----PRIRKRL--LARILEM-----FFQLOVGLIOQWVPT-IONSMKP 311  
DB 302 YOSKHTLPSINPKTFYVAVFVIALSMEFILFLHVAALSKTKAMENDTTPQISMGL 361  
QY 312 KMDYSRIERLKLAVENHILWIFFWLPHSCINAVALEMQGDEPFYRDWNSSEVTF 371  
DB 362 ENLN---II--WKLILP-----WLFRLMALDDGIDIPENNIRCD-----NNYSSL 404  
QY 372 YFQWNNIPVHKVCIRHFYKMLRRGSSKMMART--GVFLASAFHEHYVSVPLMFRPLM 429  
DB 405 AFWRAMHRSYKMKVRYIYIPL--GGSKNVLTSIAVFSVAIWHD-----IELKTL-LW 456  
QY 430 AFTGMAQIP-----LAWF-----VGRFPGANYGNAVWLSLI 462  
DB 457 GMLIVLFLFLPFIATQIFSHYTDVAVWYRHVCAVAVEN-----IWMMI 500

## RESULT 14

E91019

en-glycerol-3-phosphate permease (imported) - Escherichia coli (strain O157:H7, substrat

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004

C:Accession: E91019

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurukawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

Gawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: E91019

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-452 <HAV>

A:Cross-references: UNIPROT:O8X5A0; UNIPARC:UPI00000D0CED; GB:BA000007; PIDN:BA036548.1;

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: EC93125

C:Superfamily: hexose phosphate transport protein unpt

Query Match 4.5%; Score 117.5; DB 2; Length 452;

Best Local Similarity 18.7%; Pred. No. 0.066;

Matches 68; Conservative 52; Mismatches 120; Indels 123; Gaps 17;

143 PAVAAFOVE---KRLAVGALTEQ-----AGLLHVNATLTCFPAAVVLYVESTP 191

67 FALSGISIAVGSFKTMSVSDRNPVFLPAGLIL-----AAAVMLFMGFVP 114

192 VGSLLAAMHTLFLKLF-----SYDVNSWCRARAKAASAGKASAAA- 237

115 ----WATSSIAVMFVLFLCGFMGMPGCGRTVHWMSQKERGIYSVWNCANVGAG 170

238 -PHTVSYPDNLTLYRDLVYFLFAPLTCYELNPPRSRIRKRLIRILEMFLFTQLQVGLI 296

171 IPPLFLFGMAWFMNDWHAALYMPAFC-----AIVLALFAFAMWRDTP 212

297 QOMWVPTIONSMPKMDYSRIIR-----LKLAVPNHLIWLIFPYWLFHSCLANV 349

213 QSCGLPPIEYKNDYPD-DYNEKAEQELTAKOIFMQYVLPNKLWYI-----AIVNF 264

350 AELMOFGDREFFYRDWMSSEVTFYFQNNWNIPIVHKWCIHFYKPMLRGSSKKMARTGVFL 409

265 VYLLRYG---ILDM-----SPTYLKE-----VKHF-----ALDKSSW----- 293

410 ASAFPHEYLVSVPFLMFLMAFTGMAAOIPLAMFVGRFQGNYGNAAVWLSLIIGOPIAV 469

294 -AYFFEY-----AGIPGTLICGMSDKVFRGNRGATGVFFMTLV--TIAT 336

QY 470 LMY 472

DB 337 IVY 339

RESULT 15

G85863

en-glycerol-3-phosphate permease [imported] - Escherichia coli (strain O157:H7, substrat

C:Genetics:

A:Gene: g1pt

C:Superfamily: hexose phosphate transport protein unpt

Query Match 4.5%; Score 117.5; DB 2; Length 452;

Best Local Similarity 18.7%; Pred. No. 0.066;

Matches 68; Conservative 52; Mismatches 120; Indels 123; Gaps 17;

143 PAVAAFOVE---KRLAVGALTEQ-----AGLLHVNATLTCFPAAVVLYVESTP 191

67 FALSGISIAVGSFKTMSVSDRNPVFLPAGLIL-----AAAVMLFMGFVP 114

192 VGSLLAAMHTLFLKLF-----SYDVNSWCRARAKAASAGKASAAA- 237

115 ----WATSSIAVMFVLFLCGFMGMPGCGRTVHWMSQKERGIYSVWNCANVGAG 170

238 -PHTVSYPDNLTLYRDLVYFLFAPLTCYELNPPRSRIRKRLIRILEMFLFTQLQVGLI 296

171 IPPLFLFGMAWFMNDWHAALYMPAFC-----AIVLALFAFAMWRDTP 212

297 QOMWVPTIONSMPKMDYSRIIR-----LKLAVPNHLIWLIFPYWLFHSCLANV 349

213 QSCGLPPIEYKNDYPD-DYNEKAEQELTAKOIFMQYVLPNKLWYI-----AIVNF 264

350 AELMOFGDREFFYRDWMSSEVTFYFQNNWNIPIVHKWCIHFYKPMLRGSSKKMARTGVFL 409

265 VYLLRYG---ILDM-----SPTYLKE-----VKHF-----ALDKSSW----- 293

410 ASAFPHEYLVSVPFLMFLMAFTGMAAOIPLAMFVGRFQGNYGNAAVWLSLIIGOPIAV 469

294 -AYFFEY-----AGIPGTLICGMSDKVFRGNRGATGVFFMTLV--TIAT 336

QY 470 LMY 472

DB 337 IVY 339

Search completed: May 5, 2006, 12:34:18

Job time : 43 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2006, 12:33:51 ; Search time 47 Seconds

(without alignments)  
858.420 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MGDGSSRRRTGSRPSSHG.....VLMYVHDYVLYNTEAPAAEA 488

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/5\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/6\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/H\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/ECTUS\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/RE\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2594	100.0	488	2	US-09-165-042-1
2	2299.5	88.6	500	2	US-09-326-203A-17
3	2278.5	87.8	498	2	US-09-899-645A-8
4	2273.5	87.6	498	2	US-09-103-754A-5
5	2039	78.6	386	2	US-09-103-754A-4
6	1106	42.6	496	2	US-09-326-203A-18
7	791.5	30.1	503	2	US-09-326-203A-2
8	779.5	30.1	503	2	US-09-593-359-4
9	656.5	25.3	341	2	US-09-593-359-2
10	551	21.2	163	2	US-09-774-639-230
11	431	16.6	75	2	US-09-774-639-125
12	375.5	14.5	522	2	US-09-165-042-3
13	375.5	14.5	522	2	US-09-949-016-11030
14	370.5	14.3	409	2	US-09-326-203A-23
15	367.5	14.2	502	2	US-09-328-857A-4
16	343.5	13.2	525	2	US-09-328-857A-3
17	343	13.2	550	1	US-08-121-057-4
18	343	13.2	550	1	US-08-509-187D-4
19	343	13.2	550	4	US-09-121-396-4
20	343	13.2	550	4	PCR-UB93-09704A-4
21	337	13.0	60	2	US-09-774-639-237
22	336.5	13.0	432	2	US-09-326-203A-35
23	312	12.0	51	2	US-09-165-042-9
24	281.5	10.9	429	2	US-09-326-203A-24
25	281.5	10.9	610	2	US-09-709-457-2
26	275.5	10.6	419	2	US-09-248-796A-17982
27	259	10.0	48	2	US-09-774-639-238

28	255	9.8	47	2	US-09-774-639-239	Sequence 239, App
29	145	5.6	53	2	US-09-165-042-8	Sequence 8, Appli
30	142	5.5	53	2	US-09-165-042-7	Sequence 7, Appli
31	137	5.3	52	2	US-09-165-042-6	Sequence 6, Appli
32	134	5.2	52	2	US-09-774-639-235	Sequence 235, App
33	132	5.1	52	2	US-09-165-042-5	Sequence 5, Appli
34	118	4.5	18	2	US-09-774-639-236	Sequence 236, App
35	112	4.3	30	2	US-09-165-042-17	Sequence 17, Appl
36	111	4.3	250	2	US-09-326-203A-22	Sequence 22, Appl
37	109.5	4.2	409	2	US-09-248-796A-14863	Sequence 14863, A
38	106.5	4.1	33	2	US-09-165-042-10	Sequence 10, Appl
39	103.5	4.0	457	2	US-09-489-039A-8293	Sequence 8293, Ap
40	103	4.0	343	2	US-09-489-039A-9368	Sequence 9368, Ap
41	101.5	3.9	432	2	US-09-255-368-2	Sequence 2, Appli
42	101.5	3.9	432	2	US-09-405-558-2	Sequence 2, Appli
43	101.5	3.9	432	2	US-09-538-036-2	Sequence 2, Appli
44	101	3.9	576	2	US-09-540-236-2286	Sequence 2286, Ap
45	100.5	3.9	487	2	US-09-790-838-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1														
US-09-165-042-1														
; Sequence 1, Application US/09165042														
; Patent No. 6100077														
; GENERAL INFORMATION:														
; APPLICANT: Sturley, Stephen L.														
; APPLICANT: Oelkers, Peter														
; TITLE OF INVENTION: ISOLATION OF A GENE ENCODING DIACYLGLYCEROL														
; FILE REFERENCE: 0575/56331														
; CURRENT APPLICATION NUMBER: US/09/165,042														
; CURRENT FILING DATE: 1998-10-01														
; NUMBER OF SEQ ID NOS: 32														
; SOFTWARE: PatentIn Ver. 2.0														
; SEQ ID NO 1														
; LENGTH: 488														
; TYPE: PRT														
; ORGANISM: Yeast														
US-09-165-042-1														
Query Match														
Best Local Similarity 100.0%; Score 2594; DB 2; Length 488;														
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0														
QY	1	MGDRSSRRRTGSRPSSHGGGGPAABEEVRDAAAGPDVGAAGDAPAPAPKDGDAVG	60											
DB	1	MGDRSSRRRTGSRPSSHGGGGPAABEEVRDAAAGPDVGAAGDAPAPAPKDGDAVG	60											
QY	61	SGHWEI	CHRI	ODSIFSSDSGS	NYRGILNWCVMVLI	ISNARLFLENLIKYGILVDP	IOV	120						
DB	61	SGHWEI	CHRI	ODSIFSSDSGS	NYRGILNWCVMVLI	ISNARLFLENLIKYGILVDP	IOV	120						
QY	121	VSLFKD	HSMPAPCLVIAAVFAA	FOVERKLA	VGALTBOAGLL	LVANLATIL	CEPA	180						
DB	121	VSLFKD	HSMPAPCLVIAAVFAA	FOVERKLA	VGALTBOAGLL	LVANLATIL	CEPA	180						
QY	181	AVVL	LVESIT	TVGSLALMA	HTILFLKLF	BYRDVNS	WCRARAKASAGKASAA	PHT	240					
DB	181	AVVL	LVESIT	TVGSLALMA	HTILFLKLF	BYRDVNS	WCRARAKASAGKASAA	PHT	240					
QY	241	VSYPNL	TYRDL	YFLP	PTCYELAN	FRBRI	RKFLRL	ILEMLFT	TOLOVGLIQOM	300				
DB	241	VSYPNL	TYRDL	YFLP	PTCYELAN	FRBRI	RKFLRL	ILEMLFT	TOLOVGLIQOM	300				
QY	301	VPTI	NSMKPR	DMOYSRI	IERLLKLA	VPNHLIFYWL	FHSG	CLNVAEIMOGDRE	360					
DB	301	VPTI	NSMKPR	DMOYSRI	IERLLKLA	VPNHLIFYWL	FHSG	CLNVAEIMOGDRE	360					
QY	361	YRDW	NSES	VYTF	WQNNI	PVHK	KCIR	HFYK	PMLR	RGSSK	WMARTGV	LASFHE	YLV	420
DB	361	YRDW	NSES	VYTF	WQNNI	PVHK	KCIR	HFYK	PMLR	RGSSK	WMARTGV	LASFHE	YLV	420

Db 361 YRDWNSESVTYFMQNMNIPVHKWCI RHFYKEMLRGSSKMMARIGVFLASAFHEYLVS 420  
Qy 421 VPLRMFLMAFTGMAAQIPLAMFVGRFGNGNNAVWMLSLIIGPIATLVYHYVYAN 480  
Db 421 VPLRMFLMAFTGMAAQIPLAMFVGRFGNGNNAVWMLSLIIGPIATLVYHYVYAN 480  
Qy 481 YEAPAAEA 488  
Db 481 YEAPAAEA 488

RESULT 2  
US-09-326-203A-17  
; Sequence 17, Application US/09326203A  
; Patent No. 6444876

; GENERAL INFORMATION:  
; APPLICANT: Laesener, Mike  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; FILE REFERENCE: 17045/00/WO  
; CURRENT FILING DATE: 1999-06-04  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088,143  
; PRIOR FILING DATE: 1998-06-05  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-09-326-203A-17

Query Match 88.6%; Score 2299.5; DB 2; Length 500;  
Best Local Similarity 85.9%; Pred. No. 1.8e-225;  
Matches 427; Conservative 26; Mismatches 31; Indels 13; Gaps 3;

Qy 1 MGRD---GSSRRRTGSRPSSHGCGGPAABEEVDAAAGDPVGAADAPAPAP----- 51  
Db 1 MGRDGAAGSSRRRTGSRVSIQSGSGPVDSEEDVDAAGDPVGAADAPAPAPAH 60  
Qy 52 --NKGDAGVSGHWEHCHRLQDSLFSDSGFSNYRGILNMCVWMLISNARLFLENLKY 109  
Db 61 TRDKRQSVGSGHWEHCHRLQDSLFSDSGFSNYRGILNMCVWMLISNARLFLENLKY 120  
Qy 110 KYGDIIVDPIDVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTEQAGLLHY 169  
Db 121 KYGDIIVDPIDVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTEQAGLLHY 180  
Qy 170 ANLATITCPPAVAVLVESITPVGSLALMAHTIILFKLFSYRDVNSWC--RRARAKAS 227  
Db 181 VNLATITCPPAVAVLVESITPVGSLALMAHTIILFKLFSYRDVNSWC--RRARAKAS 240  
Qy 228 AGKKAASAAAPHVSYPDNLTYRDLYFLFAPTLCEYLNFPRSPRIKRFLLRRIEMLF 287  
Db 241 AGKKAASAAAPHVSYPDNLTYRDLYFLFAPTLCEYLNFPRSPRIKRFLLRRIEMLF 300  
Qy 288 FTQLOVGLIQOMNVPTIONSMPKPKMDYSRIIERLLKLAVPNHLIWLIFFYWLFHSCIN 347  
Db 301 FTQLOVGLIQOMNVPTIONSMPKPKMDYSRIIERLLKLAVPNHLIWLIFFYWLFHSCIN 360  
Qy 348 AVAELMOGDFREYVDWNSESVTYFMQNMNIPVHKWCI RHFYKEMLRGSSKMMARIGV 407  
Db 361 AVAELMOGDFREYVDWNSESVTYFMQNMNIPVHKWCI RHFYKEMLRGSSKMMARIGV 420  
Qy 408 FLASAFHEYLVS VPLRMFLMAFTGMAAQIPLAMFVGRFGNGNNAVWMLSLIIGPI 467  
Db 421 FLASAFHEYLVS VPLRMFLMAFTGMAAQIPLAMFVGRFGNGNNAVWMLSLIIGPI 480  
Qy 468 AVLMTVDHYVYLVYVYAP 484  
Db 481 AVLMTVDHYVYLVYVYAP 484

Db 481 AVLMTVDHYVYLVYVYAP 497

RESULT 3  
US-09-899-645A-8  
; Sequence 8, Application US/09899645A  
; Patent No. 694170

; GENERAL INFORMATION:  
; APPLICANT: Li, Chun Ping  
; APPLICANT: Zheng, Peizhong  
; APPLICANT: Nicholas, Scott  
; TITLE OF INVENTION: METHODS FOR REGULATING BETA-OXIDATION IN PLANTS  
; FILE REFERENCE: 35718/235742  
; CURRENT FILING DATE: 2001-07-05  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 498  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-899-645A-8

Query Match 87.8%; Score 2278.5; DB 2; Length 498;  
Best Local Similarity 85.1%; Pred. No. 2.4e-223;  
Matches 421; Conservative 29; Mismatches 34; Indels 11; Gaps 3;

Qy 1 MGRD---GSSRRRTGSRPSSHGCGGPAABEEVDAAAGDPVGAADAPAPAP----- 51  
Db 1 MGRDGAAGSSRRRTGSRVSIQSGSGPVDSEEDVDAAGDPVGAADAPAPAPAH 60  
Qy 52 NKGDAGVSGHWEHCHRLQDSLFSDSGFSNYRGILNMCVWMLISNARLFLENLKY 111  
Db 61 TRDKRQSVGSGHWEHCHRLQDSLFSDSGFSNYRGILNMCVWMLISNARLFLENLKY 120  
Qy 112 GILVDPIDVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTEQAGLLHYAN 171  
Db 121 GILVDPIDVSLFLKDPHSPAPCLVIANVFAVAAFQVEKRLAVGALTEQAGLLHYAN 180  
Qy 172 LATITCPPAVAVLVESITPVGSLALMAHTIILFKLFSYRDVNSWC--RRARAKAS 229  
Db 181 LATITCPPAVAVLVESITPVGSLALMAHTIILFKLFSYRDVNSWC--RRARAKAS 240  
Qy 230 QLOVGLIQOMNVPTIONSMPKPKMDYSRIIERLLKLAVPNHLIWLIFFYWLFHSCINAV 349  
Db 241 QLOVGLIQOMNVPTIONSMPKPKMDYSRIIERLLKLAVPNHLIWLIFFYWLFHSCINAV 360  
Qy 350 AELMOGDFREYVDWNSESVTYFMQNMNIPVHKWCI RHFYKEMLRGSSKMMARIGV 409  
Db 361 AELMOGDFREYVDWNSESVTYFMQNMNIPVHKWCI RHFYKEMLRGSSKMMARIGV 420  
Qy 410 ASAFHEYLVS VPLRMFLMAFTGMAAQIPLAMFVGRFGNGNNAVWMLSLIIGPIAV 469  
Db 421 ASAFHEYLVS VPLRMFLMAFTGMAAQIPLAMFVGRFGNGNNAVWMLSLIIGPIAV 480  
Qy 470 LMYVHDYVYLVYVYAP 484  
Db 481 LMYVHDYVYLVYVYAP 495

RESULT 4  
US-09-103-754A-5  
; Sequence 5, Application US/09103754A  
; Patent No. 6344548  
; GENERAL INFORMATION:  
; APPLICANT: Farese, Robert  
; APPLICANT: Cases, Sylvaine

APPLICANT: Smith, Steven  
APPLICANT: Erickson, Sandra  
TITLE OF INVENTION: Diacylglycerol O-acyltran  
TITLE OF INVENTION: sfetase  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Bozicevic & Reed  
STREET: 285 Hamilton Avenue, Suite 200  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/103,754A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 6510-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650 327 3400  
TELEFAX: 650 327 3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 498 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-103-754A-5

Query Match 87.6%; Score 2273.5; DB 2; Length 498;  
Best Local Similarity 84.8%; Pred. No. 7.8e-223;  
Matches 420; Conservative 29; Mismatches 35; Indels 11; Gaps 3;  
QY 1 MGBR---GSSRRRRRGGSSHHGGGPPAAEEVDDAAAGPVGAAGDAPAPAP----- 51  
DB 1 MGBRGAGSSSRRRRGGSSHHVVGSGSGPKVEEDVDAVSPDLGAGDAPAPAPAPAPATR 60  
QY 52 NKDAGAGSGHWEIACHRLDLSFSSDGSFNSYGIINWCVMLILSNARLFLENLKY 111  
DB 61 DKGDTVGVGGWIDRCHRLDLSFSSDGSFNSYGIINWCVMLILSNARLFLENLKY 120  
QY 112 GILVDPIDVSVFLKDPHSPWAPCLVIANVFAVAFQVEKRLAVGALTQAGLILHYAN 171  
DB 121 GILVDPIDVSVFLKDPYSWAPPCVITASNIFVAAFOIEKRLAVGALTQAGLILHYAN 180  
QY 172 LATIICFPAAVVLVESITPVGSLIAMAHTILFKLFSYRDVNSWC--RRARAAASAG 229  
DB 181 LATIICFPAAVVALVESITPVGSLIAMAHTILFKLFSYRDVNSWCORRKAARAVSTG 240  
QY 230 KKASAAAPHVTSYSDNLTYYRDLVYFLFAPTLCYELNPRSPRIKRLFLRLILEMLPT 289  
DB 241 KKVSSAAAOQAVSYSDNLTYYRDLVYFLFAPTLCYELNPRSPRIKRLFLRLILEMLPT 300  
QY 290 QLVGLIQOMWVPTIIONSMKPKMDYSRIERLKLAVPNHLIWLIFPYWLFHSCLANV 349  
DB 301 QLVGLIQOMWVPTIIONSMKPKMDYSRIERLKLAVPNHLIWLIFPYWLFHSCLANV 360  
QY 350 AELMOPGREFRDMWNSSEVTYFQNMNIPVHKACIRHFYPMIRGSSKMATGVPFL 409  
DB 361 AELMOPGREFRDMWNSSEVTYFQNMNIPVHKACIRHFYPMIRGSSKMATGVPFL 420

QY 410 ASAFHEYLVSVPRLMFLMAFTGMAAOIPLAMFYGRFFQNGYNAAVLSLIGOPILAV 469  
DB 421 TSAPFHEILVSVPRLMFLMAFTGMAAOIPLAMFYGRFFQNGYNAAVLSLIGOPILAV 480  
QY 470 LMYVDYVLYNYP 484  
DB 481 LMYVDYVLYNYP 495

RESULT 5  
US-09-103-754A-4  
Sequence 4, Application US/09103754A  
Patent No. 6344548  
GENERAL INFORMATION:  
APPLICANT: Farese, Robert  
APPLICANT: Cases, Sylvaine  
APPLICANT: Smith, Steven  
APPLICANT: Erickson, Sandra  
TITLE OF INVENTION: Diacylglycerol O-acyltran  
TITLE OF INVENTION: sfetase  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Bozicevic & Reed  
STREET: 285 Hamilton Avenue, Suite 200  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/103,754A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 6510-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650 327 3400  
TELEFAX: 650 327 3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 386 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-103-754A-4

Query Match 78.6%; Score 2039; DB 2; Length 386;  
Best Local Similarity 99.5%; Pred. No. 4.3e-199;  
Matches 384; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 103 LFLNLIRYGLVNDPIQVSVFLKDPHSPWAPCLVIANVFAVAFQVEKRLAVGALTQEO 162  
DB 1 LFLNLIRYGLVNDPIQVSVFLKDPHSPWAPCLVIANVFAVAFQVEKRLAVGALTQEO 60  
QY 163 AGLILHVNANLITICFPAAVVLVESITPVGSLIAMAHTILFKLFSYRDVNSWCRRAR 222  
DB 61 AGLILHVNANLITICFPAAVVLVESITPVGSLIAMAHTILFKLFSYRDVNSWCRRAR 120  
QY 223 AKASAGKASSAAAPHVTSYSDNLTYYRDLVYFLFAPTLCYELNPRSPRIKRLFLRLI 282  
DB 121 AKASAGKASSAAAPHVTSYSDNLTYYRDLVYFLFAPTLCYELNPRSPRIKRLFLRLI 180

QY 283 LEMHFFTOLOVGLIQOMWVPTIONSMPKPKOMDYSRRIERLKLAVPNHLIWLIFFWLF 342  
DB 181 LEMHFFTOLOVGLIQOMWVPTIONSMPKPKOMDYSRRIERLKLAVPNHLIWLIFFWLF 240  
QY 343 HSCINAAVLELMOFGREFYRDWMNSESVTYFWQNNIPVHKCIRHFKKIPMLRRSSKXM 402  
DB 241 HSCINAAVLELMOFGREFYRDWMNSESVTYFWQNNIPVHKCIRHFKKIPMLRRSSKXM 300  
QY 403 ARTGVFLASAFHEHLYVSVPLRMFLMAFTGMAQIPLAMFVGRFPOGNGAAYVLSLI 462  
DB 301 ARTGVFLASAFHEHLYVSVPLRMFLMAFTGMAQIPLAMFVGRFPOGNGAAYVLSLI 360  
QY 463 IGOPIAVLMYVHDYVLYVLYEAPAAEA 488  
DB 361 IGOPIAVLMYVHDYVLYVLYEAPAAEA 386

RESULT 6  
US-09-326-203A-18  
Sequence 18, Application US/09326203A  
Patent No. 644876  
GENERAL INFORMATION:  
APPLICANT: Laesener, Mike  
APPLICANT: Ruzinskiy, Diane  
TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
FILE REFERENCE: 17045/00/WO  
CURRENT APPLICATION NUMBER: US/09/326,203A  
CURRENT FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: 60/088,143  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/108,389  
PRIOR FILING DATE: 1998-11-12  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 18  
LENGTH: 496  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-326-203A-18

Query Match 42.6%; Score 1106; DB 2; Length 496;  
Best Local Similarity 43.7%; Pred. No. 7,8e-104;  
Matches 220; Conservative 89; Mismatches 155; Indels 40; Gaps 6;  
QY 6 SRRRTGSRPSHGGCGGPAABEEVRDAAGDPVGAAGDAPAPPNKDGCGVSGSME 65  
DB 5 TGRRRR---QPSSTNGSLASSRR-----SSFAQNGSSRKSSSEKRGFCE 46  
QY 66 LRCHRLQDLSFSSDSGFNRYGILNMCVMLILSNARLFLNLIKYGLIIVPPIQVVSFL 125  
DB 47 KVVHTAQDSLSTSGMTNFRGFPLSLILLVLSNGRALEEVIVKXGILITPLQWISITFV 106  
QY 126 KDPH---SWPAPCLVIANVFAVAAFOVEKRLAVALTEQAGLLHVNALATILCEPAAY 182  
DB 107 EHHYSIMSPMLATILGNITQILSVFGMEKILERGMVGFAVAYTSLVIMHTIPVVV 166  
QY 163 VLVVESTIPVGSILAMHTILFLKFSYRDVNSCRBARAKAASGKASSA----- 236  
DB 167 TLTHKMKPLMSVVMGVIVLEALKFISYGHVNYARDRRKITLKTQVTDLAKKTCDP 226  
QY 237 -----APHTVS--YPDNLTYRDLVYFLFAFTLCYELNFPSPRIKRFLRLRIE 284  
DB 227 KQPMOLKDELMSHQMAAQYPAULITSNITYFPAATLTCEPEKFPRLIRIKHFLIKRTYE 286  
QY 285 MLFPTOLOVGLIQOMWVPTIONSMPKPKOMDYSRRIERLKLAVPNHLIWLIFFWLHS 344  
DB 287 LIFLSFLIAALVQGVVVPVTRNSMKPLSEMEYSRCLEBLLKLAIDNHILIMLFFYTFHS 346  
QY 345 CLNAAELMOFGREFYRDWMNSESVTYFWQNNIPVHKCIRHFKKIPMLRRSSKXMAR 404  
DB 347 FLNLIAELLRADREFYRDWMNSESVTYFWQNNIPVHKCIRHFKKIPMLRRSSKXAF 406

QY 405 TGVFLASAFHEHLYVSVPLRMFLMAFTGMAQIPLAMFVGRFPOGNGAAYVLSLI 463  
DB 407 FVVFVSAFHEHLYVSVPLRMFLMAFTGMAQIPLAMFVGRFPOGNGAAYVLSLI 466  
QY 464 GQPIAVLMYVHDYVLYVLYEAPAAEA 487  
DB 467 GQPIAVLMYVHDYVLYVLYEAPAAEA 490

RESULT 7  
US-09-326-203A-2  
Sequence 2, Application US/09326203A  
Patent No. 644876  
GENERAL INFORMATION:  
APPLICANT: Laesener, Mike  
APPLICANT: Ruzinskiy, Diane  
TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
FILE REFERENCE: 17045/00/WO  
CURRENT APPLICATION NUMBER: US/09/326,203A  
CURRENT FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: 60/088,143  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/108,389  
PRIOR FILING DATE: 1998-11-12  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-326-203A-2

Query Match 30.5%; Score 791.5; DB 2; Length 520;  
Best Local Similarity 38.0%; Pred. No. 9,5e-72;  
Matches 194; Conservative 77; Mismatches 174; Indels 65; Gaps 17;  
QY 8 RRRRTGSRPSHGGCGGPAABEEVRDAAGDPVGAAGDAPAPPNKDGCGVSGSME 48  
DB 25 RLRRKRSRSDSSNGLLSGDNNSPSDVGAAPVDRIDSVNDDAQG-TANLAGDNG 83  
QY 49 PAPNKDAGVSGSGHWEIIRC-----HRLQDLSFSSDSGF-SYRIRILNMCVVM 96  
DB 84 GGDNNGGGCGGEGRGNADATFYRPSVPAHRRARESPSSDAIFKSHAGLFNLCVVL 143  
QY 97 ILNARLFLNLIKYGLIIV-DPIQVSLFLKDPHSPAPCLVIANVFAVAAFOVEKRLA 155  
DB 144 IAVNSRLIENLMKYLIRDPFSSRLRD---WPLFCCISLSTFPLAFTVEKVL 200  
QY 156 VQALTEQAGLLHVNALATILCFPAAYVVLVESITPVGSLALAMHTILFLKFSYR-- 212  
DB 201 QKIYSEPVVILHIIITWTEVLVYVYTLRCDSAFLSGVTLMILT-CIVMLKLSYATS 259  
QY 213 -DVNSWCRARAASAKASAKASSAAAPHTVSYPDNLTTRDYVYFLFAFTLCYELNFPSP 271  
DB 260 YDIRS-----LANADKAMP-----EVS--YVSLKSLAYPMVAVTLCYQSPSPSA 304  
QY 272 RIRKRFILRLIEMHFFTOLOVGLIQOMWVPTIONSMPKPKOMDYSRRIERLKLAVPN 330  
DB 305 CIRKGMVARQPAKIVIFGFGFIIIEQYINIVANSKPLKGDLLXA--IRVLTLSVFN 362  
QY 331 HLIMLIFFWLHSCINAAVLELMOFGREFYRDWMNSESVTYFWQNNIPVHKCIRHFK 390  
DB 363 LYVLMCMFYCEPHLMINILAEILCEGDEFEYKDMWNKASVGDYRMNMMPVHKMVRHLY 422  
QY 391 KPMLRSSKXMMATGVFLASAFHEHLYVSVPLRMFLMAFTGMAQIPLAMFVGRFPOG 450  
DB 423 FPCIRSKLPKTIATIIIFLVAVHELCIAVBCFLKLMALGIMFOVPLV-FTTNYLOE 481  
QY 451 NY-----GNAAYV-LSLIIGOPIAVLMYVHD 475  
DB 482 RFGSTVGMMITWFIPLIFGQPMCVLLIYHD 511

```
RESULT 8
US-09-593-359-4
; Sequence 4, Application US/09593359
; Patent No. 6552250
; GENERAL INFORMATION:
; APPLICANT: Laroche, Andre J.
; APPLICANT: Nykiforuk, Cory L.
; APPLICANT: Weselake, Randall J.
; TITLE OF INVENTION: Diacylglycerol O-acyltransferase
; FILE REFERENCE: 24015U50
; CURRENT APPLICATION NUMBER: US/09/593,359
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Brassica napus
; OTHER INFORMATION: DGAT1
US-09-593-359-4

Query Match      30.1%; Score 779.5; DB 2; Length 503;
Best Local Similarity 38.7%; Pred. No. 1.5e-70;
Matches 191; Conservative 83; Mismatches 168; Indels 51; Gaps 15;

QY 10 RRTGSRPSHG-----GGGPAABEEVDAAGPDV--GAADGAPAPAKKGDGCV 59
DB 26 RKKSSDSSNGLLSPTSPSDVGAABAEEDRVDAABEBAOQTANLAGDAETRESAGG- 84
QY 60 GSGHWELRC-----HRLQDSLFSDDSGF--SNYRGILNWCVMILISNARLFLENLI 109
DB 85 -----DVAFTYRSPVPAHRTRESPLSSDAIFKOSHAGLFNLCVYVAVNVRLLIENLM 139
QY 110 KYGILV-DPIQVVSJFLKDPHWPAPCLVIANVAVAAFOVEKRLAVGALTQOAGLLIH 168
DB 140 KYGMLIRDFWFSSTSLRD---WPLFMCCLSISVFPPLAFTVEKAVLQKFISEPVAILIH 196
QY 169 VANLATICPPAAVVLVESTIPVGSLLAMNHTLFLKLFPSYRVNWSKRAABAKASA 228
DB 197 VITMTVLPYVYVTLRCDASFLSGVTMLLT--CIWMLKLVSAHTS--YDIRFLANSA 252
QY 229 GKKAASAAPIHTVSYPDNLTYRDLTYFLFAPTLCEYLNFPSPRIKRFLLRIILEMFLF 288
DB 253 DK-----VDPEISY--YVSLKSLAYFMVAAPTLCQPSYSPSCIRKGMVARQLAKYIF 304
QY 289 TQLOVGLIQQWNPPTIQNSMKPKF--DMYDSRIIEELKLAVPNHLIWLIFPYWLFHSCIN 347
DB 305 TGLMGFIIEQYINPIVRNSKHPLKGDLLYA--IERVLKLSVBNLYVWLCMFYCFHLMIN 362
QY 348 AVAELMORGEDEEYRDMMNSSESVTYFMQNNINPVYKMKCIRHFYKMLERGSSKMMARCTGV 407
DB 363 TLAEELCGDEEFYKDMNNAKSVGDYWRMNNPVYKMMVRRHYVFFCLIKIPKPAVAILIA 422
QY 408 FLASAFHEBYLVSVPLRMFLMAFTGMMAOIPLAFVGRFPGQNGNAAVWL-----SLI 462
DB 423 FLVGAFFHELCIAVPCRLFNMAFMGIMFQVPLV-FITNFIQERFGSVGNMIRGSASCI 481
QY 463 IGPPIAVLMVYVD 475
DB 482 FGQPMCGILYYHD 494

RESULT 9
US-09-593-359-2
; Sequence 2, Application US/09593359
; Patent No. 6552250
; GENERAL INFORMATION:
; APPLICANT: Laroche, Andre J.
; APPLICANT: Nykiforuk, Cory L.
; APPLICANT: Weselake, Randall J.
; TITLE OF INVENTION: Diacylglycerol O-acyltransferase
; FILE REFERENCE: 24015U50
```

```
; CURRENT APPLICATION NUMBER: US/09/593,359
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Brassica napus
; OTHER INFORMATION: DGAT2
US-09-593-359-2
```

```
Query Match      25.3%; Score 656.5; DB 2; Length 341;
Best Local Similarity 41.9%; Pred. No. 3e-58;
Matches 144; Conservative 62; Mismatches 117; Indels 21; Gaps 8;
```

```
QY 138 IANVFAVAFOVEKRLAVGALTQOAGLLHYANLATICPPAAVVLVESTIPVGSLLA 197
DB 4 LLSIFPLAFAFVEELVLOKCISEBPVIFLHYITMTVLPYVYVTLRCDASFSLGDTLM 63
QY 198 LMAHTILFLKLFESYRDVNSWCRARAKAAGKASAAAPHTVSYPDNLTYRDLTYFLF 257
DB 64 LIT-CIVMLKLVSYAHTYDIR-----TIANSSDKANP-----EVSY--YVSLKSLAYFNL 111
QY 258 APTLCYELNFPSPRIKRFLLRIILEMFLFTOLQVGLIQQWNPPTIQNSMKPKF--DMY 316
DB 112 APTLCYQPSYRSPCIRKGMVARQPAKLVIFGLMGFIIEQYINPIVRNSKHPLKGDLLY 171
QY 317 SRIIRLKLAVPNHLIWLIFPYWLFHSCLNAAVAELMOPGDEEYRDMMNSSESVTYFMON 376
DB 172 A--IERVLKLSVBNLYVWLCMFYCFHLMINLAEELCGDEEFYKDMNNAKSVGDYWRM 229
QY 377 WNIPIHAKCIRHFYKPMLERGSSKMMARCTGVFLASAFHEBYLVSVPLRMFLMAFTGMM 436
DB 230 WNIPIHAKMVRKRVYFPCIRIKIPKVPAILIIFLSAVFHELCIAVPCRLFNMAFMGIMF 289
QY 437 QIPLAMFVGRFPGQNGNAAVWL-----SLIIGPIAVLMVYVD 475
DB 290 QVPLV-FITNFIQERFGSVGNMIRGSASCIFGQPMCGILYYHD 332
```

```
RESULT 10
US-09-774-639-230
; Sequence 230, Application US/09774639
; Patent No. 6806351
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 230
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-774-639-230
```

```
Query Match      21.2%; Score 551; DB 2; Length 163;
Best Local Similarity 89.0%; Pred. No. 5.7e-48;
Matches 105; Conservative 2; Mismatches 5; Indels 6; Gaps 1;
```

```
QY 377 WNIPIVH-----KMCIRHFYKPMLERGSSKMMARCTGVFLASAFHEBYLVSVPLRMFLMA 430
DB 46 WMAVNLIRFPCCPLCRHFYKPMLERGSSKMMARCTGVFLASAFHEBYLVSVPLRMFLMA 105
QY 431 FTGMMAOIPLAMFVGRFPGQNGNAAVWLSLIIGPIAVLMVYVDYVLANEAPAA 488
DB 106 FTGMMAOIPLAMFVGRFPGQNGNAAVWLSLIIGPIAVLMVYVDYVLANEAPAA 163
```

RESULT 11  
US-09-774-639-125  
Sequence 125 Application US/09774639  
Patent No. 6606351  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 90 Human Secreted Proteins  
FILE REFERENCE: P2013P1  
CURRENT APPLICATION NUMBER: US/09/774,639  
CURRENT FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04  
NUMBER OF SEQ ID NOS: 371  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 125  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (75)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-774-639-125

Query Match 16.6%; Score 431; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 3,1e-36;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 314 MDYSRIIRLLKLVNHLIMLIFYYWLFHSCINAAVAELMOFGDREFRDMWNSSTVYF 373  
DB 1 MDYSRIIRLLKLVNHLIMLIFYYWLFHSCINAAVAELMOFGDREFRDMWNSSTVYF 60  
QY 374 WNNNIPVHKWCIR 387  
DB 61 WNNNIPVHKWCIR 74

RESULT 12  
US-09-165-042-3  
Sequence 3 Application US/09165042  
Patent No. 610077  
GENERAL INFORMATION:  
APPLICANT: Sturley, Stephen L.  
APPLICANT: Oelkers, Peter  
TITLE OF INVENTION: ISOLATION OF A GENE ENCODING DIACYLGLYCEROL  
FILE REFERENCE: 0575/56331  
CURRENT APPLICATION NUMBER: US/09/165,042  
CURRENT FILING DATE: 1998-10-01  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 3  
LENGTH: 522  
TYPE: PRT  
ORGANISM: Yeast  
US-09-165-042-3

Query Match 14.5%; Score 375.5; DB 2; Length 522;  
Best Local Similarity 26.3%; Pred. No. 2,4e-29;  
Matches 127; Conservative 75; Mismatches 189; Indels 91; Gaps 20;  
QY 45 DAPAPAP-----NKGDAGVSGHWEILRCHRLQDSLFSDSGFSNRYGILNW-----CVV 94  
DB 79 DKPLPPPPGSLSTRQPSLG-----KQKVFIRKSLDLDELMEVGHFRITIMFIAGLCVF 134  
QY 95 ML-----ILSNARLFE-NLIKYGILVDPQVVS---LFLKDPHSPAPCLVIAANVFA 144  
DB 135 IISTLAIDFIDEGRLLEFDLILFSFQGLPLALVTWVPMFLST-----L 178  
QY 145 VAAQVEKRLAVGALTEQAGL-LHVNANLITLCFPAVVLLVESITPVGSLALMAHTI 203  
DB 179 LAPYQALRMARGTWTATGIGCALLAAHAAVLLC-ALPVAHVAEHQLPAPASRCVLAPEQV 237

QY 204 LFL-KLFSY-RDVNSMCRARAKAASAGKASSAAAPHTVSPDNLTYRDLVYLFAPTL 261  
DB 238 RFLMKSYFLAEAVPGTLRAR-----RGEIGQAFSFSY-----LYFLFCPTL 280  
QY 262 CYELNPPSPRIIRKQFLRLRLLEMLFFPQLOVGIQQMMVPTION-SMKPKMDYSRII 320  
DB 281 IYRETYPTPTPVNRYAKNFAQALGCVLACFLGLGCVFVFAVMSREPEST--RALV 337  
QY 321 ERLKLVNHLIMLIFYYWLFHSCINAAVAELMOFGDREFRDMWNSSTVYFWQNNIP 380  
DB 338 LSIHATLPGFEMLLIFPAFLHGMNLNFAEMLRFGDMRFYRDMWNSSTVYRTNIVV 397  
QY 381 VHKWCIRHFKYKMLR--RGSSKMMARTGVPLASAFHEHYLVSPLRMR-----LMAFTG 433  
DB 398 VHDWLSYVQDGLRLGAPARAVAMLGVLVSVAHEYIFCFVLGFFYPVMLILFLVIG 457  
QY 434 MMAQIPL-----AWVGRPFQGNYNAAVMSLIIGQPIAVLMYHDYVLYNVEPAA 486  
DB 458 GMLFMNHDQRTGPAW-----NVLMTWMLFLGQIGVSLYQCEWYARRH-CPLP 505  
QY 487 EA 488  
DB 506 QA 507

RESULT 13  
US-09-949-016-11030  
Sequence 11030 Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11030  
LENGTH: 522  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-11030

Query Match 14.5%; Score 375.5; DB 2; Length 522;  
Best Local Similarity 26.3%; Pred. No. 2,4e-29;  
Matches 127; Conservative 75; Mismatches 189; Indels 91; Gaps 20;  
QY 45 DAPAPAP-----NKGDAGVSGHWEILRCHRLQDSLFSDSGFSNRYGILNW-----CVV 94  
DB 79 DKPLPPPPGSLSTRQPSLG-----KQKVFIRKSLDLDELMEVGHFRITIMFIAGLCVF 134  
QY 95 ML-----ILSNARLFE-NLIKYGILVDPQVVS---LFLKDPHSPAPCLVIAANVFA 144  
DB 135 IISTLAIDFIDEGRLLEFDLILFSFQGLPLALVTWVPMFLST-----L 178  
QY 145 VAAQVEKRLAVGALTEQAGL-LHVNANLITLCFPAVVLLVESITPVGSLALMAHTI 203  
DB 179 LAPYQALRMARGTWTATGIGCALLAAHAAVLLC-ALPVAHVAEHQLPAPASRCVLAPEQV 237  
QY 204 LFL-KLFSY-RDVNSMCRARAKAASAGKASSAAAPHTVSPDNLTYRDLVYLFAPTL 261  
DB 238 RFLMKSYFLAEAVPGTLRAR-----RGEIGQAFSFSY-----LYFLFCPTL 280  
QY 262 CYELNPPSPRIIRKQFLRLRLLEMLFFPQLOVGIQQMMVPTION-SMKPKMDYSRII 320  
DB 281 IYRETYPTPTPVNRYAKNFAQALGCVLACFLGLGCVFVFAVMSREPEST--RALV 337

321 BRLKAVPNHILWIFPYWLFHSCINAAVAELMOFGDREFYRDMNNSSVTYFMQNNIP 380  
338 LSLHATIPGIMLILIPAFILHCHWLNFAELMRGDRMYRDMNNSISFSNYRTMNV 397  
391 VHKWICIRHFYKEMLR--RGSSKMMARATGVFLASAFHEHYLVSPLRMFR-----LMAFTG 433  
398 VHDMLYSYVQDGLHLLGARARGVAMLGVFVLSAVAHEXIIFCFVIGFFYPVWLLIFLVIG 457  
434 MMHQIPL-----AMFGRFPGQNYGNAAVWLSLIIGPIAVLMTYHDYVLYLNEAPAA 486  
458 GMLNFMWMDORTGPAW-----NVLMTMLFLGGIGVSLYCEQWYARRH-CPUP 505  
487 EA 488  
506 QA 507

RESULT 14  
US-09-326-203A-23  
Sequence 23, Application US/09326203A  
Patent No. 6444876  
GENERAL INFORMATION:  
APPLICANT: Laesener, Mike  
APPLICANT: Ruzinskiy, Diane  
TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
FILE REFERENCE: 17045/00/WO  
CURRENT APPLICATION NUMBER: US/09/326,203A  
CURRENT FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: 60/088,143  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/108,389  
PRIOR FILING DATE: 1998-11-12  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 23  
LENGTH: 409  
TYPE: PRT  
ORGANISM: murine  
PUBLICATION INFORMATION:  
JOURNAL: J. Biol. Chem.  
VOLUME: 270  
PAGES: 26192-26201  
DATE: 1995  
US-09-326-203A-23

Query Match 14.3%; Score 370.5; DB 2; Length 409;  
Best Local Similarity 27.5%; Pred. No. 5, 5e-29;  
Matches 123; Conservative 69; Mismatches 164; Indels 91; Gaps 18;

71 LQBSIFSSDGSFNSYRGILNMCVWML-----SNARLLE-NLIRYGLVDP 118  
4 LIDELFEVD---HIRTYYMFIALILFVLSITVVDYDEGRVLEFLLAVAGKEPPT 59  
119 QV-----VSLFKDPHSPAPCLVIAAVFAAFOVEKRLAVGALLTQAGLL 166  
60 VITWMMWMLSTLSIPYFLPOP-W-----AHGYSKSHPLLYSLV-----HGLL 102  
167 LHVNLATILCPRAVVLIVESTIPVGSLLAMATILFLKLSYRDVNSWCRARAKAA 226  
103 FLVPLQLS-VLGVFPVYVLAATLPASRPILILBQIRILMK-----ASHFRENIPRYL 155  
227 SAGKASAAAPAHYTSYDNLTYRDLVYFLFAPLTCYELNFRSPRIKRFILIRILEM- 285  
156 NAAEKSSKDPILPVN-----QYLYFLFAPTLIYRDNVBRTPVWGVAAQFLOVF 207  
286 --LFTQVGLIQOMVPTION-SMKPKMDYSRIIRLKLAV-----PNHILWILF 337  
208 GCLFVYVY---IFERLCAPLFENIKQBPSS-----ARVAVLCVNSILPGVILFLS 256  
338 FYWLFHSCINAAVAELMOFGDREFYRDMNNSSVTYFMQNNIPVHKWICIRHFYKEMLRG 397

257 FFAFLHCHWLNFAEMLRFGDRMFYKDMNNSISYSNYRTMNVVHDMWLYVYVYKDLLMWF 316  
398 SSK--MARTGVFLASAFHEHYLVPLR-----NFRIMAFNGMAQIPLAMFVRPQG 450  
317 SKRFSAAMLAVFALSAVVEHALAICLSYPVLFVLFMEFGM-----AFNFIYVDSRRK 372  
451 NYGNAAVWLSLIIGPIAVLMTYHDY 477  
373 PIMNIMWVASLFLGIGLILCFYSQEMY 399

RESULT 15  
US-09-328-857A-4  
Sequence 4, Application US/09328857A  
Patent No. 6579974  
GENERAL INFORMATION:  
APPLICANT: Caeses, Sylvaine  
APPLICANT: Farese, Robert  
APPLICANT: No. 6579974ak, Sabine  
APPLICANT: Erickson, Sandra  
TITLE OF INVENTION: No. 6579974el Acyl CoA:Cholesterol Acyl  
FILE REFERENCE: 6510-104US1  
CURRENT APPLICATION NUMBER: US/09/328,857A  
CURRENT FILING DATE: 1999-06-08  
PRIOR APPLICATION NUMBER: 60/090,354  
PRIOR FILING DATE: 1998-06-23  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 502  
TYPE: PRT  
ORGANISM: human  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (191)...(191)  
OTHER INFORMATION: Xaa is any amino acid  
US-09-328-857A-4

Query Match 14.2%; Score 367.5; DB 2; Length 502;  
Best Local Similarity 26.7%; Pred. No. 1, 5e-28;  
Matches 127; Conservative 64; Mismatches 187; Indels 97; Gaps 20;

45 DAPAPAKDDAGVGSCHWELCRQLDSIFSSDGSFNSYRGILNMCVWML-----IL 98  
79 DKPLPFP---PPGSLSELMEXQHRTTYHMF-----LGLCVFIITSLAIDPI 124  
99 SNARLLE---NLIRYGLVDP 152  
125 DEGRLLLEFDDPLIFSGQL--PLALVTWVPMFLST-----LLAVYQALR 166  
153 RLAVGALLTEQGL--LHVNLATILCPRAVVLIVESTIPVGSLLAMATILFL-KLF 209  
167 LMARGTWTOAGTGCALLAAHAAVXCALPVVA--VEHQLPASRCVIVFEQVREFLMKSY 224  
210 SY-RDVNSWCRARAKASAKKASAAAPHTVSYPDLTYRDLVYFLFAPLTCYELNFP 268  
225 SPLREAVPGTIRAR-----RGGIOAPSESS-----LYFLFCTPLIYRETYP 267  
269 RSPRIKRFILIRILEMFLTQVGLIQOMVPTION-SMKPKMDYSRIIRLKLKLA 327  
268 RTPYRMMYVAKNFMQALGCVLYACFLIGRLCVPFAMMSDPST---RALVLSILHAT 324  
328 VPHNLWILFYWLFHSCINAAVAELMOFGDREFYRDMNNSSVTYFMQNNIPVHKWICIR 387  
325 LPGLFMILLIFPAFLHCHWLNFAEMLRFGDRMFYKDMNNSISFSNYRTMNVVHDMWYS 384  
368 HPYKEMLR--RGSSKMMARATGVFLASAFHEHYLVSPLRMFR-----LMAFTGMAAQIPL 440  
365 YVQDGLHLLGARARGVAMLGVFVLSAVAHEXIIFCFVIGFFYPVWLLIFLVIGGMLNPM 444  
441 -----AMFGRFPGQNYGNAAVWLSLIIGPIAVLMTYHDYVLYLNEAPAAEA 488

Db 445 HDQRTGPAW-----NVLMTWTLFLGGGIQVGLYCOEWYARRH-CPLPQA 487

Search completed: May 5, 2006, 12:35:10  
Job time : 49 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2006, 12:44:55 ; Search time 28 Seconds  
(without alignments)  
806.677 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MGDGRSSRRRTGSRPSSHG.....YLMYVHDYVLYNEAPAAEA 488

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA New:  
1: /SIDS5/ptcodata/1/pubppaa/US08\_NEW\_PUB.pep1.\*  
2: /SIDS5/ptcodata/1/pubppaa/US06\_NEW\_PUB.pep.\*  
3: /SIDS5/ptcodata/1/pubppaa/US07\_NEW\_PUB.pep.\*  
4: /SIDS5/ptcodata/1/pubppaa/US08\_NEW\_PUB.pep.\*  
5: /SIDS5/ptcodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*  
6: /SIDS5/ptcodata/1/pubppaa/US09\_NEW\_PUB.pep.\*  
7: /SIDS5/ptcodata/1/pubppaa/US10\_NEW\_PUB.pep.\*  
8: /SIDS5/ptcodata/1/pubppaa/US10\_NEW\_PUB.pep1.\*  
9: /SIDS5/ptcodata/1/pubppaa/US11\_NEW\_PUB.pep.\*  
10: /SIDS5/ptcodata/1/pubppaa/US11\_NEW\_PUB.pep.\*  
11: /SIDS5/ptcodata/1/pubppaa/US11\_NEW\_PUB.pep1.\*  
12: /SIDS5/ptcodata/1/pubppaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2594	100.0	488	US-11-317-983-5	Sequence 5, Appl1
2	2278.5	87.8	498	US-11-317-983-4	Sequence 4, Appl1
3	791.5	30.5	520	US-11-317-983-2	Sequence 2, Appl1
4	779.5	30.1	503	US-11-317-983-9	Sequence 9, Appl1
5	778	30.0	547	US-11-317-983-25	Sequence 25, Appl1
6	656.5	25.3	341	US-11-317-983-8	Sequence 8, Appl1
7	551	21.2	163	US-10-986-501-298	Sequence 98, Appl1
8	431	16.6	74	US-10-986-501-155	Sequence 155, App
9	375.5	14.5	522	US-10-995-561-1030	Sequence 1020, Ap
10	337	13.0	60	US-10-986-501-305	Sequence 305, App
11	259	10.0	48	US-10-986-501-306	Sequence 306, App
12	255	9.8	47	US-10-986-501-307	Sequence 307, App
13	134	5.2	26	US-10-986-501-303	Sequence 303, App
14	118	4.5	9	US-10-986-501-304	Sequence 304, App
15	114.5	4.4	484	US-10-467-657-3820	Sequence 1082, Ap
16	103	4.0	868	US-10-821-234-1082	Sequence 2, Appl1
17	101.5	3.9	432	US-10-992-577-2	Sequence 37, Appl1
18	101.5	3.9	432	US-11-223-294-37	Sequence 6, Appl1
19	100.5	3.9	487	US-11-040-595-6	Sequence 20418, A
20	100	3.9	314	US-11-096-568A-20418	Sequence 302, App
21	98	3.8	17	US-10-986-501-302	

22	91.5	3.5	423	11	US-11-079-463-8361	Sequence 8361, Ap
23	91.5	3.5	4097	9	US-10-501-035-263	Sequence 263, App
24	91	3.5	449	11	US-11-079-463-5270	Sequence 5270, Ap
25	90.5	3.5	482	11	US-11-072-512-3794	Sequence 3794, Ap
26	90	3.5	673	9	US-10-055-877-321	Sequence 321, App
27	89.5	3.5	615	9	US-10-821-234-873	Sequence 873, App
28	89	3.4	369	9	US-10-506-454-749	Sequence 749, App
29	89	3.4	516	11	US-11-188-298-20291	Sequence 20291, A
30	88.5	3.4	4128	9	US-10-770-726-77	Sequence 77, Appl1
31	88	3.4	381	11	US-11-096-568A-10876	Sequence 10876, A
32	88	3.4	400	9	US-10-454-437-306	Sequence 306, App
33	88	3.4	423	9	US-10-454-437-304	Sequence 304, App
34	88	3.4	461	11	US-11-079-463-9682	Sequence 9682, Ap
35	87	3.4	831	7	US-09-941-095-5	Sequence 5, Appl1
36	87	3.4	831	11	US-11-198-746-5	Sequence 5, Appl1
37	87	3.4	831	11	US-11-198-794-5	Sequence 5, Appl1
38	87	3.4	831	11	US-11-183-211-5	Sequence 5, Appl1
39	87	3.4	831	11	US-11-198-657-5	Sequence 1811, A
40	86.5	3.3	486	11	US-11-188-298-13811	Sequence 9, Appl1
41	86.5	3.3	2589	11	US-11-216-660-9	Sequence 35, Appl1
42	86	3.3	330	11	US-11-110-274-4	Sequence 37, Appl1
43	86	3.3	330	11	US-11-040-218-35	Sequence 37, Appl1
44	86	3.3	344	11	US-11-040-218-37	Sequence 39, Appl1
45	86	3.3	346	11	US-11-040-218-39	

#### ALIGNMENTS

```
RESULT 1
US-11-317-983-5
; Sequence 5, Application US/11317983
; Publication No. US20060090222A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David C
; APPLICANT: Wei, Yangdou
; APPLICANT: Jako, Collette C
; TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants
; FILE REFERENCE: 3015-5684,105
; CURRENT APPLICATION NUMBER: US/11/317,983
; CURRENT FILING DATE: 2005-12-23
; PRIOR APPLICATION NUMBER: US/11/117,005
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 09/623,514
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: PCT/CA99/01202
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/112,812
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 488
; TYPE: PRT
; ORGANISM: human
US-11-317-983-5

Query Match      100.0%; Score 2594; DB 10; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.2e-224;
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGRSSRRRTGSRPSSHGGAEEYRDAAAGDVGAAADGAPAPAPKDDAGVG 60
   |||
DB 1 MGDGRSSRRRTGSRPSSHGGAEEYRDAAAGDVGAAADGAPAPAPKDDAGVG 60
   |||
QY 61 SGHWEILRCHRIQDSIFSSDSGFSNRYGILNVCVWLISNARLFLENLTKYGIIVDP1QV 120
   |||
DB 61 SGHWEILRCHRIQDSIFSSDSGFSNRYGILNVCVWLISNARLFLENLTKYGIIVDP1QV 120
   |||
QY 121 VSLFLKDPHSPAPCLVTIAANVFAVAFOVEKRLAVGALTTEOAGILLHVANLATTLCFPA 180
   |||
DB 121 VSLFLKDPHSPAPCLVTIAANVFAVAFOVEKRLAVGALTTEOAGILLHVANLATTLCFPA 180
   |||
```

```

Qy 181 AVTLVESITPVGSLALMAHTILFLKLFPSYRDVNSMCRBARAKASAKKASAAAPHT 240
    |||
Db 181 AVTLVESITPVGSLALMAHTILFLKLFPSYRDVNSMCRBARAKASAKKASAAAPHT 240
Qy 241 VSYPDNLTYRDLTYFLFAPTLCEYELNPPSPRIKRFLLRILEMLFTQLQVGLIQQM 300
    |||
Db 241 VSYPDNLTYRDLTYFLFAPTLCEYELNPPSPRIKRFLLRILEMLFTQLQVGLIQQM 300
Qy 301 VPTIOMSKPDKMDYSRIERLLKLAVPNHLIMLFFYWLPHSCLNAVAELMOGDEPF 360
    |||
Db 301 VPTIOMSKPDKMDYSRIERLLKLAVPNHLIMLFFYWLPHSCLNAVAELMOGDEPF 360
Qy 361 YRDWNSSEVTFWQNNMIPVHKCIRHFYKPMLRSSKMMARTGVFLASAFPHYLYS 420
    |||
Db 361 YRDWNSSEVTFWQNNMIPVHKCIRHFYKPMLRSSKMMARTGVFLASAFPHYLYS 420
Qy 421 VPLRMFLMAFTGMAAQIPLAMFVGRFQGNYGNAAVWLSLIIGQPIAVLWVHDYVYL 480
    |||
Db 421 VPLRMFLMAFTGMAAQIPLAMFVGRFQGNYGNAAVWLSLIIGQPIAVLWVHDYVYL 480
Qy 481 YEAPAAEA 488
    |||
Db 481 YEAPAAEA 488

```

## RESULT 2

```

US-11-317-983-4
; Sequence 4, Application US/11317983
; Publication No. US2006009022A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David C
; APPLICANT: Wei, Yangdou
; APPLICANT: Jakó, Coleete C
; TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants
; FILE REFERENCE: 3015-5684, IUS
; CURRENT APPLICATION NUMBER: US/11/317,983
; PRIOR FILING DATE: 2005-12-23
; PRIOR APPLICATION NUMBER: US/11/117,005
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 09/623,514
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: PCT/CA99/01202
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/112,812
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 498
; TYPE: PRT
; ORGANISM: mouse
US-11-317-983-4

```

Query Match 87.8%; Score 2278.5; DB 10; Length 498;  
 Best Local Similarity 85.1%; Pred. No. 8.8e-196;  
 Matches 421; Conservative 29; Mismatches 34; Indels 11; Gaps 3;

```

Qy 1 MGDR--GSSRRRTGSSPSSHGCGPAAAEEDVDAAGPVGAAADAPAP----- 51
    |||
Db 1 MGDRGAGSSRRRTGSSPSSHGCGPAAAEEDVDAAGPVGAAADAPAPAPATR 60
Qy 52 NKDDGAGVSGHWEIHCRLDLSFSSDGSFENYGIIMCVMLILSNARLFLENLKY 111
    |||
Db 61 DKDGTSTVGDGWDLCRHLQDLSFSSDGSFENYGIIMCVMLILSNARLFLENLKY 120
Qy 112 GILVDPIDVSLFLKDPHSPAPCLVIANVPAVAFOVEKRLAVGALTEQAGLLHYAN 171
    |||
Db 121 GILVDPIDVSLFLKDPHSPAPCLVIANVPAVAFOVEKRLAVGALTEQAGLLHYAN 180
Qy 172 LATIICPPAAVAVLVESITPVGSLALMAHTILFLKLFPSYRDVNSMCRBARAKASAG 229
    |||

```

```

Db 181 LATIICPPAAVAVLVESITPVGSLALMAHTILFLKLFPSYRDVNSMCRBARAKASAG 240
Qy 220 KKASAAAPHTVSPDNLTYRDLTYFLFAPTLCEYELNPPSPRIKRFLLRILEMLFT 289
    |||
Db 241 KKVSGAAAOQAVSPDNLTYRDLTYFLFAPTLCEYELNPPSPRIKRFLLRILEMLFT 300
Qy 290 QLQVGLIQQMWVPTIOMSKPDKMDYSRIERLLKLAVPNHLIMLFFYWLPHSCLNAV 349
    |||
Db 301 QLQVGLIQQMWVPTIOMSKPDKMDYSRIERLLKLAVPNHLIMLFFYWLPHSCLNAV 360
Qy 360 AELMOGDEPFYRDWNSSEVTFWQNNMIPVHKCIRHFYKPMLRSSKMMARTGVFL 409
    |||
Db 361 AELMOGDEPFYRDWNSSEVTFWQNNMIPVHKCIRHFYKPMLRSSKMMARTGVFL 420
Qy 420 ASAFPHYLYSVPLRMFLMAFTGMAAQIPLAMFVGRFQGNYGNAAVWLSLIIGQPIAV 469
    |||
Db 421 TSAPFHELYSVPLRMFLMAFTGMAAQIPLAMFVGRFQGNYGNAAVWLSLIIGQPIAV 480
Qy 470 LMYVDYVYLVNYPEAP 484
    |||
Db 481 LMYVDYVYLVNYPEAP 495

```

## RESULT 3

```

US-11-317-983-2
; Sequence 2, Application US/11317983
; Publication No. US2006009022A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David C
; APPLICANT: Wei, Yangdou
; APPLICANT: Jakó, Coleete C
; TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants
; FILE REFERENCE: 3015-5684, IUS
; CURRENT APPLICATION NUMBER: US/11/317,983
; PRIOR FILING DATE: 2005-12-23
; PRIOR APPLICATION NUMBER: US/11/117,005
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 09/623,514
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: PCT/CA99/01202
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/112,812
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-317-983-2

```

Query Match 30.5%; Score 791.5; DB 10; Length 520;  
 Best Local Similarity 38.0%; Pred. No. 8.7e-63;  
 Matches 194; Conservative 77; Mismatches 174; Indels 65; Gaps 17;

```

Qy 8 RRRRTGSSPSSHGCGPAAAEEDVDAAGPVGAAADAPAP----- 48
    |||
Db 25 RLRKRKSSDSNGLLSGDMNSPDDVGAAPADVRDIDSVNDDAG--TANLAQDNG 83
Qy 49 PAPKDGAGVSGHWEIHCRLDLSFSSDGSFENYGIIMCVMLILSNARLFLENLKY 96
    |||
Db 84 GGDNNGGGRRGGEGGNADATFTYRPSVPAHRRARRESPLSDAIFKOSHAGLFNLCVVL 143
Qy 97 ILSNARLFLENLKYGLIV-DPIQVSLFLKDPHSPAPCLVIANVPAVAFOVEKRLA 155
    |||
Db 144 IAVNSRLIENLMRYGWLIRTFDFWSSRSRLD---WPLFMCCISLSTFPLAAFTVEKVL 200
Qy 156 VGALTEQAGLLHYANLATIICPPAAVAVLVESITPVGSLALMAHTILFLKLFPSYR--- 212
    |||
Db 201 QKTISEPVYIFLHIITMTVELVPEVYVTLRCDSAFLSGSVTLMULT-CIVMLKTVSYARTS 259
Qy 213 -DVNSMCRBARAKASAKKASAAAPHTVSPDNLTYRDLTYFLFAPTLCEYELNPPSP 271

```

```

Db      260 YDIRS-----LANAADKANP-----EVSY--YVSLKSLAVFMVAPTLCPQSPYRGA 304
Qy      272 RIRKFLRLRIEMLEFFQOLQVGLIQOMNVPTIOMSKPKF-DMDYSRIERBLKLAVERN 330
Db      305 CIRKMWVARQPKLVIPTGFMGFIIEQYINPVRNSKHLKGDLLYA--IERVLKLSVFN 362
Qy      331 HLHILIFPFWMLPSCINAVAEIMQFGRREFFYRDMMNSESUTYFWQWNI PVHKMCIRHPY 390
Db      363 LYVWLCMYCFPHMLNLIABELLCFGRREFFYRDMMNKSVDGYWRMNNPVHKMVRHY 422
Qy      391 KEMLRGSSKMMARTGVFLASAFPEHYLVSVPLRMFRMAFTGMAQIPLAMFVGAFPG 450
Db      423 FPCLSKSLPKTLAIITIAFLVASVFBELCIAVPCRLFKLWAFGIMFQVPLV-FITVYQGE 481
Qy      451 NY-----GNAAVW-LSLITQPIAVLMYVAD 475
Db      482 RFGSTVGNNIFWFIICIFGQPMCVLLYYHD 511

```

## RESULT 4

```

US-11-317-983-9
; Sequence 9, Application US/11317983
; Publication No. US2006090222A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David C
; APPLICANT: Wei, Yangdou
; APPLICANT: Jako, Colette C
; TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants
; FILE REFERENCE: 3015-5684.1US
; CURRENT APPLICATION NUMBER: US/11/317,983
; PRIOR FILING DATE: 2005-12-23
; PRIOR APPLICATION NUMBER: US/11/117,005
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 09/623,514
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: PCT/CA99/01202
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/112,812
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Brassica napus
US-11-317-983-9

```

Query Match 30.1%; Score 779.5; DB 10; Length 503;

Best Local Similarity 38.7%; Pred. No. 9,8e-62;

Matches 191; Conservative 83; Mismatches 168; Indels 51; Gaps 15;

```

Qy      10 RRTGSPSSHG-----GGPAAAEERVDAAAGPDV-GAAGDAPAPAKDGDAGV 59
Db      26 RKRSSSDSNGLSTSPSDVGAABERDRDAABEAGTANLAGDADRSGAG- 84
Qy      60 GSGHWELRC-----HRLDSLFSSDSGF-SYRGLMKCVMLISNARLLENLI 109
Db      85 ----DVFRTYRPSVPAHRRRTRESPLSDAIFKOSHAGLFNLCVVVLAIVANSRLIENIM 139
Qy      110 KYGILV-DPIQVSLFLKDPHMPAPCLVIANVAVAVAFOVERKRLAVALTEQAGLLH 168
Db      140 KYGMLIRIDPWFSSISLRD--WPLFMCCLSLVFPLAFTVEKVLQKFTISEPVAILIH 196
Qy      169 VAKLATTILCFPAAVVLVESITPVGSLALMAHTILFLKPSYRDVNSWCRBARAKASA 228
Db      197 VITWTEVLYPYVYTLRCDASAFLSGVTMLMLT-CIVWLKLVSAHTS--YDIRLTAASA 252
Qy      229 GKXSSAAAPRTVSPDULTRDLYYFLPAFTLCYELNPPSPRIKRFILRIEMLEPF 288
Db      253 DK-----VDPEISY--YVSLKSLAVFMVAPTLCPQSPYRSPCIRKMWVARQKLAVIF 304

```

```

Qy      289 TOLQVGLIQOMNVPTIOMSKPKF-DMDYSRIERBLKLAVERNHLIWLIFPWLPHSCIN 347
Db      305 TGLMGPFIIEQYINPVRNSKHLKGDLLYA--IERVLKLSVFNLYVWLCMFCEPHMLN 362
Qy      348 AVALIEMQFGRREFFYRDMMNSESUTYFWQWNI PVHKMCIRHPYKEMLRGSSKMMARTGV 407
Db      363 ILAELLCGDRREFFYRDMMNKSVDGYWRMNNPVHKMVRHYFPCLRIKIRKVPAILIA 422
Qy      408 FLASAFPEHYLVSVPLRMFRMAFTGMAQIPLAMFVGAFPGQNGNAVWL-----SLI 462
Db      423 FLVASVFBELCIAVPCRLFKLWAFGIMFQVPLV-FITNLOERFGSSVGNMIFGSASCI 481
Qy      463 ICGPIAVLMYVAD 475
Db      482 FQGPWCGLLYYHD 494

```

## RESULT 5

```

US-11-317-983-25
; Sequence 25, Application US/11317983
; Publication No. US2006090222A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David C
; APPLICANT: Wei, Yangdou
; APPLICANT: Jako, Colette C
; TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants
; FILE REFERENCE: 3015-5684.1US
; CURRENT APPLICATION NUMBER: US/11/317,983
; PRIOR FILING DATE: 2005-12-23
; PRIOR APPLICATION NUMBER: US/11/117,005
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 09/623,514
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: PCT/CA99/01202
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/112,812
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-317-983-25

```

Query Match 30.0%; Score 778; DB 10; Length 547;

Best Local Similarity 36.1%; Pred. No. 1.5e-61;

Matches 194; Conservative 77; Mismatches 174; Indels 92; Gaps 17;

```

Qy      8 RRRRTGSPSSHG-----GGPAAAEER-----VRDAAAGPDVGAAGDAPA 48
Db      25 RLRRKRSKSDSNGLLSGSDNNSPSDVGAAPADVRDIDSVVNDAGG-TANLAGDNNG 83
Qy      49 PAPKPDGAGVSGHWELRC-----HRLDSLFSSDSGF----- 82
Db      84 GGDNNGGARGGEBRGMDATFTYRPSVPAHRRARESPLSDAIFKOSHAGLFNLCVVL 143
Qy      83 ----SNYRGLMKCVMLISNARLLENLIKYGILV-DPIQVSLFLKDP 128
Db      144 IAVNSRLIENIMKSHAGLFNLCVVVLAIVANSRLIENIMKYGWILFRDWFSSRSLRD- 202
Qy      129 HSWPAPCLVIANVAVAVAFOVERKRLAVALTEQAGLLHYVANLATTILCFPAAVVLVES 188
Db      203 --WPLFMCCLSLVFPLAFTVEKVLQKFTISEPVAILIITWTEVLYPYVYTLRCD 260
Qy      189 ITPVGSLLALMAHTILFLKPSYR-----DVNSWCRBARAKASAGKXSSAAAPRTVGY 244
Db      261 AFLSGVTMLMLT-CIVWLKLVSAHTSYDIRS-----LANAADKANP-----EVSY- 305
Qy      245 DNLTYRDLYYFLPAFTLCYELNPPSPRIKRFILRIEMLEPFQOLQVGLIQOMNVPTI 304
Db      306 --YVSLKSLAVFMVAPTLCPQSPYRSPCIRKMWVARQPAKLVIFPTGFMGFIIEQYINP 364

```

```
Qy 305 QNSMKPFK-DMDYSRIERLLKLAVERNHLIMLFFPYWLFSHSCNAVAELMORGDEPRD 363
Db 365 RNSKHEPLGDLLEYA--IERVLKLSVFNLYVLCMFYCFEHLMLNLAEILCTGDEDFK 422
Qy 364 WNSSEVTVFVQNNNI PVHKVCIRHFYKPMRLRGSSKMMARTGVFLASAFHEHYLSVPL 423
Db 423 WNNKSVGDYVRMMNMPVHKMMVRIITFPCLRSKIPKTLAIITAVLSAVFHELCIAVPC 482
Qy 424 RMFRIMATGMAOIPFLAMFVGRFQGNV---GNAAVW-LSLIGQPIAVLMTYVD 475
Db 483 RLFLKMAELGIMFQVPLV-FITNVLQERFGVSGNMIMFIFCIGQPCVCLLYHD 538
```

## RESULT 6

```
US-11-317-983-8
; Sequence 8, Application US/11317983
; Publication No. US20060090222A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David C
; APPLICANT: Wei, Yangdou
; APPLICANT: Jako, Colette C
; TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants
; FILE REFERENCE: 3015-5684.1US
; CURRENT APPLICATION NUMBER: US/11/317,983
; PRIOR APPLICATION NUMBER: US/11/117,005
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 09/623,514
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: PCT/CA99/01202
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/112,812
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Brassica napus
US-11-317-983-8
```

```
Query Match 25.3%; Score 656.5; DB 10; Length 341;
Best Local Similarity 41.9%; Pred. No. 66-51;
Matches 144; Conservative 62; Mismatches 117; Indels 21; Gaps 8;
```

```
Qy 138 IAAVFAVAATQVERKRLAVGALTQAGLLHVAANLATTLCFPAAVLLVESTTPVGSLLA 197
Db 4 LLSISIFLAATFVELVLOKICISEPVVIFLHVITMTFVLYPVVYTLCDSAFLSGDTLM 63
Qy 198 LMAHTILFLKLEPSYDVNWSKCRARAKASAGKSSAAAPRTVYPNDLTTRDYLYPLF 257
Db 64 LLT-CIVLKVSVANTYDIR---TLANSSDKANP-----EVSY--YVSLKSLAYFML 111
Qy 258 APTLCYELNPPSPRIKRFLLRIEMLFTQLQVGLIQGMVNTIIONSMKPFK-DMDY 316
Db 112 APTLCYOSYRSRPSIRKGMVARQAKVITFGLMGFIIEQYINIVANSKRLPKGDLXY 171
Qy 317 SRIIRLLKLAVERNHLIMLFFPYWLFSHSCNAVAELMORGDEPRDWMNSSEVTVFQ 376
Db 172 A--IERVLKLSVFNLYVLCMFYCFEHLMLNLAEILCTGDEDFYKDMNNAKSVGDYVRM 229
Qy 377 WNIPIHKVCIRHFYKPMRLRGSSKMMARTGVFLASAFHEHYLSVPLMFRIMATGMA 436
Db 230 WNNPVMKMMVRRHVFPCLRIKIPKVPALITIAVLSAVFHELCIAVPCFLNLMAGTME 289
Qy 437 QIPLAMFVGRFQGNVGNAAVWL-----SLIGQPIAVLMTYVD 475
Db 290 QVPLV-FITNVLQERFGVSGNMIMGTSASCFIGQPCVCLLYHD 332
```

## RESULT 7

```
US-10-986-501-298
; Sequence 298, Application US/10986501
; Publication No. US20050244845A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P2C1
; CURRENT APPLICATION NUMBER: US/10/986,501
; CURRENT FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US/10/621,363
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/969,730
; PRIOR FILING DATE: 2001-10-06
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: PCT/US98/16235
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/056,371
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,732
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,366
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,364
; PRIOR FILING DATE: 1997-08-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 298
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-986-501-298
```

```
Query Match 21.2%; Score 551; DB 9; Length 163;
Best Local Similarity 89.0%; Pred. No. 6,56-42;
Matches 105; Conservative 2; Mismatches 5; Indels 6; Gaps 1;
```

```
Qy 377 WNIPIH-----KVCIRHFYKPMRLRGSSKMMARTGVFLASAFHEHYLSVPLMFRIMA 430
Db 46 WNNVNLNRRFPCLRIHFYKPMRLRGSSKMMARTGVFLASAFHEHYLSVPLMFRIMA 105
Qy 431 FTGMAOIPFLAMFVGRFQGNVGNAAVWLSLIIGQPIAVLMTYVDYVLTNEAPAAEA 488
Db 106 FTGMAOIPFLAMFVGRFQGNVGNAAVWLSLIIGQPIAVLMTYVDYVLTNEAPAAEA 163
```

## RESULT 8

```
US-10-986-501-155
; Sequence 155, Application US/10986501
; Publication No. US20050244845A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P2C1
; CURRENT APPLICATION NUMBER: US/10/986,501
; CURRENT FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US/10/621,363
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/969,730
; PRIOR FILING DATE: 2001-10-06
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: PCT/US98/16235
; PRIOR FILING DATE: 1998-08-04
```

;; PRIOR APPLICATION NUMBER: 60/056,371  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: 60/056,732  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: 60/056,366  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: 60/056,364  
;; PRIOR FILING DATE: 1997-08-19  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 373  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 155  
;; LENGTH: 74  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-986-501-155

Query Match 16.6%; Score 431; DB 9; Length 74;  
Best Local Similarity 100.0%; Pred. No. 1.3e-31;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 MDYRIIRLKLAVPNHLIWIIFYYWLFHSCNAVAELMFGDREFYRDMMNSSEVTYF 373  
Db 1 MDYSRIIRLKLAVPNHLIWIIFYYWLFHSCNAVAELMFGDREFYRDMMNSSEVTYF 60  
Qy 374 WQNNIPVHKWCIR 387  
Db 61 WQNNIPVHKWCIR 74

RESULT 9  
US-10-995-561-1030  
;; Sequence 1030, Application US/10995561  
;; Publication No. US20050272054A1  
;; GENERAL INFORMATION:  
;; APPLICANT: CARGILL, Michele et al.  
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
;; TITLE OF INVENTION: DETECTION AND USES THEREOF  
;; FILE REFERENCE: C001559  
;; CURRENT APPLICATION NUMBER: US/10/995,561  
;; CURRENT FILING DATE: 2004-11-24  
;; NUMBER OF SEQ. ID NOS: 85702  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1030  
;; LENGTH: 522  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-995-561-1030

Query Match 14.5%; Score 375.5; DB 9; Length 522;  
Best Local Similarity 26.3%; Pred. No. 1.4e-25;  
Matches 127; Conservative 75; Mismatches 189; Indels 91; Gaps 20;

Qy 45 DAPAPAP-----NKGDAGVSGHMEIRLQDLSFSDGFSNYGILNM-----CVV 94  
Db 79 DKLPPEPPPSLSRTQPSLG----KQVFIRKSLDELMVEGHFRITIMFTAGLCVF 134  
Qy 95 ML-----LISNAFLIE-NIKYGIIVDPQVVS---LFLKDPHSPAPCLVIANVFA 144  
Db 135 IISTLAIDFIDEGRLLEFDLILFSPQLPLATVTPWPMFIST-----L 178  
Qy 145 VAAFOVEKRLAVGLTEQAGI-LHVNANLITLCFPAVVLVSITPVGSLAMHTI 203  
Db 179 LAPQALRLMARGTWTATGTCALLAAHAAVLC-ALPVHAVERHQLPPASRCVLVEQV 237  
Qy 204 LFL-KLFSY-RDVSMCRARAKAASAGKASAAAPHTVSYPNLTLYRDLVYFLFAPTL 261  
Db 238 RFLMKSYSPLEAHPGLTAR-----RGEIGQASFSY-----LYLFCPTL 280  
Qy 262 CYELNPRSPRIKRFILRIILELFTQLOVGLIQOMVPTTON-SMKPKMDYSRII 320  
Db 281 IYRETYRTPVYRWNVYAKNFAQALGCVLYACFTILGRLCVVPANMSREPFST---DALV 337

Qy 321 ERLKLAVPNHLIWIIFYYWLFHSCNAVAELMFGDREFYRDMMNSSEVTYFQNNIP 380  
Db 338 LSLIHAATLPGLFMILLIFPAFLHQMNAFAEMLFGRBMFRRDMNNSFSFNYYRTNNV 397  
Qy 381 VHKWCIRHFKYPMIR--RGSSKMMARFGVFLASAFHEHYLVSPILNFR-----LMAFTG 433  
Db 398 VHDWLYSYVQDGRILGARARGVAMGVFLVSAVAHEHYICFVYLGFFYPVMLILFLVIG 457  
Qy 434 MMAQIPL-----AMVEGRFGQNYGNAAVWLSLIIGQPIAVLMYVHDVYLVNAPAA 486  
Db 458 GMLNFMHMDORTGPAM-----NLMTMTPLFGQGIQVSLYCOEYARRH-CLLP 505  
Qy 487 EA 488  
Db 506 QA 507

RESULT 10  
US-10-986-501-305  
;; Sequence 305, Application US/10986501  
;; Publication No. US20050244845A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruben et al.  
;; TITLE OF INVENTION: 90 Human Secreted Proteins  
;; FILE REFERENCE: P2013P2C1  
;; CURRENT APPLICATION NUMBER: US/10/986,501  
;; CURRENT FILING DATE: 2004-11-12  
;; PRIOR APPLICATION NUMBER: US/10/621,363  
;; PRIOR FILING DATE: 2003-07-18  
;; PRIOR APPLICATION NUMBER: 09/969,730  
;; PRIOR FILING DATE: 2001-10-06  
;; PRIOR APPLICATION NUMBER: 09/774,639  
;; PRIOR FILING DATE: 2001-02-01  
;; PRIOR APPLICATION NUMBER: 60/238,291  
;; PRIOR FILING DATE: 2000-10-06  
;; PRIOR APPLICATION NUMBER: 09/244,112  
;; PRIOR FILING DATE: 1999-02-04  
;; PRIOR APPLICATION NUMBER: PCT/US98/16235  
;; PRIOR FILING DATE: 1998-08-04  
;; PRIOR APPLICATION NUMBER: 60/056,371  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: 60/056,732  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: 60/056,366  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: 60/056,364  
;; PRIOR FILING DATE: 1997-08-19  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 373  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 305  
;; LENGTH: 60  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-986-501-305

Query Match 13.0%; Score 337; DB 9; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.5e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 PKMDYSRIIRLKLAVPNHLIWIIFYYWLFHSCNAVAELMFGDREFYRDMMNSSE 369  
Db 1 PKMDYSRIIRLKLAVPNHLIWIIFYYWLFHSCNAVAELMFGDREFYRDMMNSSE 60

RESULT 11  
US-10-986-501-306  
;; Sequence 306, Application US/10986501  
;; Publication No. US20050244845A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruben et al.  
;; TITLE OF INVENTION: 90 Human Secreted Proteins

```
FILE REFERENCE: P2013P2C1
CURRENT APPLICATION NUMBER: US/10/986,501
CURRENT FILING DATE: 2004-11-12
PRIOR APPLICATION NUMBER: US/10/621,363
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/969,730
PRIOR FILING DATE: 2001-10-06
PRIOR APPLICATION NUMBER: 09/774,639
PRIOR FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: 60/238,291
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 09/244,112
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: PCT/US98/16235
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/056,371
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,732
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,366
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,364
PRIOR FILING DATE: 1997-08-19
Remaining Prior Application data removed - See File Wrapper or PALM.
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 306
LENGTH: 48
TYPE: PRT
ORGANISM: Homo sapiens
US-10-986-501-306

Query Match
Best Local Similarity 10.0%; Score 259; DB 9; Length 48;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 387 RHFKPMLRRGSSCKMARTGVFLASAPFHEHYLVSPILMFMAFTGM 434
Db 1 RHFKPMLRRGSSCKMARTGVFLASAPFHEHYLVSPILMFMAFTGM 48

RESULT 12
US-10-986-501-307
Sequence 307, Application US/10986501
Publication No. US20050244845A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 90 Human Secreted Proteins
FILE REFERENCE: P2013P2C1
CURRENT APPLICATION NUMBER: US/10/986,501
CURRENT FILING DATE: 2004-11-12
PRIOR APPLICATION NUMBER: US/10/621,363
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/969,730
PRIOR FILING DATE: 2001-10-06
PRIOR APPLICATION NUMBER: 09/774,639
PRIOR FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: 60/238,291
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 09/244,112
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: PCT/US98/16235
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/056,371
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,732
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,366
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,364
PRIOR FILING DATE: 1997-08-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 373
```

```
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 307
LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
US-10-986-501-307

Query Match
Best Local Similarity 9.8%; Score 255; DB 9; Length 47;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 435 MAQIPLAWFVGRFGQNGYNAAVWLSLITGPIAVLMTYHDYLVNT 481
Db 1 MAQIPLAWFVGRFGQNGYNAAVWLSLITGPIAVLMTYHDYLVNT 47

RESULT 13
US-10-986-501-303
Sequence 303, Application US/10986501
Publication No. US20050244845A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 90 Human Secreted Proteins
FILE REFERENCE: P2013P2C1
CURRENT APPLICATION NUMBER: US/10/986,501
CURRENT FILING DATE: 2004-11-12
PRIOR APPLICATION NUMBER: US/10/621,363
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/969,730
PRIOR FILING DATE: 2001-10-06
PRIOR APPLICATION NUMBER: 09/774,639
PRIOR FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: 60/238,291
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 09/244,112
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: PCT/US98/16235
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/056,371
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,732
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,366
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,364
PRIOR FILING DATE: 1997-08-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 373
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 303
LENGTH: 26
TYPE: PRT
ORGANISM: Homo sapiens
US-10-986-501-303

Query Match
Best Local Similarity 5.2%; Score 134; DB 9; Length 26;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 284 EMLEFDTQVGLIQOMVVPITQNSMK 309
Db 1 EMLEFDTQVGLIQOMVVPITQNSMK 26

RESULT 14
US-10-986-501-304
Sequence 304, Application US/10986501
Publication No. US20050244845A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 90 Human Secreted Proteins
FILE REFERENCE: P2013P2C1
CURRENT APPLICATION NUMBER: US/10/986,501
```

```
/ CURRENT FILING DATE: 2004-11-12
/ PRIOR APPLICATION NUMBER: US/10/621,363
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: 09/969,730
/ PRIOR FILING DATE: 2001-10-06
/ PRIOR APPLICATION NUMBER: 09/774,639
/ PRIOR FILING DATE: 2001-02-01
/ PRIOR APPLICATION NUMBER: 60/238,291
/ PRIOR FILING DATE: 2000-10-06
/ PRIOR APPLICATION NUMBER: 09/244,112
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: PCT/US98/16235
/ PRIOR FILING DATE: 1998-08-04
/ PRIOR APPLICATION NUMBER: 60/056,371
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,732
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,366
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,364
/ PRIOR FILING DATE: 1997-08-19
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 373
/ SOFTWARE: Patencin Ver. 2.0
/ SEQ ID NO 304
/ LENGTH: 18
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-986-501-304
```

```
Query Match          4.5%; Score 118; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      370 VTTFQNNNIPVHKWCIR 387
DB      1 VTTFQNNNIPVHKWCIR 18
```

```
RESULT 15
US-10-467-657-3820
/ Sequence 3820, Application US/10467657
/ Publication No. US20050260581A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqMan99, version 1.04
/ SEQ ID NO 3820
/ LENGTH: 484
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3820
```

```
Query Match          4.4%; Score 114.5; DB 9; Length 484;
Best Local Similarity 19.2%; Pred. No. 0.027;
Matches 92; Conservative 56; Mismatches 152; Indels 179; Gaps 22;
```

```
QY      131 WPAFLVIAANFEAV-----AAFOYKRLAVGALTEQGLLHVANLATIICPPAA 181
DB      5 WFMPLSVFALFFIVPLPIYGFAPKYPVONLL--LAAGMGWLXHSIPV-----FAHI 57
QY      182 VLVVESTTPVGSLLAL-----MAHTTLFLKLFPSYRDVNSWCRARAKAASA 228
```

```
DB      58 IVLVSSCVLLGELLRSDRESTRFRMLGGGIAASITVLGFYKYPD---FFRBLIAQYAGK 114
QY      229 GKXASSAABHTVSPDNLTYRDLVYFLF---AP-----TLCELNK----- 267
DB      115 G-GAIDIMPLGLSY---YTFQSVAVLYVCPPAPHAARFGMBELHLHSFFPTVTSPII 170
QY      268 -----PRSPRIKRFLRLILEMLFTQLQVGLI-QQMMVPT 303
DB      171 RAAAFKSTDEQAGALAQIRTRPRSP--VRPALAVSLILGLAKKMWLAGILAENNVSPV 229
QY      304 ION-----SKKPYQMD-YSRIT--ERLKLAVPNHLIWLIFPYWLFH 343
DB      230 FENPTQFDGSGVLAGVYGYTFQLFLDFSGYSDLVIGMMLLGFRLPKN-----FS 279
QY      344 SCLNVAELMQGDREFYRDMMNSSVYTFQNNNIPVHKWCIRRFYKPMRLRGSSKXMA 403
DB      280 APLRAA-----NIRAFWKHHSISTWTRDYIYPL--GSSKRGFL 318
QY      404 RT-----GVFLSAFFHEY-----LVSVPLRMFLMAFTGMAQIPLAW 442
DB      319 RTQNLMAAMVLSGIMHGYGNFLLWGLHGTALALMLNTGDRYFGRDALCRKLXLAFLSW 378
QY      443 FV-----GRFFQNGYNAAVW-----LSLIIGPIAVLMTYVDY 476
DB      379 FVTFHFVCLSFYVFENTANPDDAGAVFSALEFANAGGMNAFORADMILLASFASIMLTPY 437
```

```
Search completed: May 5, 2006, 12:46:20
Job time : 29 secs
```

**THIS PAGE BLANK (USPTO)**





```
OY 253 TTTTyrPheLeuPheAlaProThrLeuCySTyrGluLeuAsnPheProArgSerProArg 272
Db 3 TACTTCTTCTCTTGGCCCCCACCCTTGCTGCTACGAGCTCAACTTTCGCCGCTCCCGCC 62
OY 273 TLeaGlyAsnGlyPheLeuLeuArgArgIleLeuGluMetLeuPhePheThrGluLeuGln 292
Db 63 ATCCCGAAGCGCTTCTGCTGCGACGGATCTTGAGATCCTGTTCTTCAACCAGCTCCAG 122
OY 293 VALGlyLeuIleGlnGlnTrpMetValProThrIleGlnAsnSerMetLysProPheLys 312
Db 123 GTGGGGCTGATCCAGCAGTGAATGATGCCACCATCAAGACTTCATGAAAGCCCTTCAG 182
OY 313 AspMetAspTyrSerArgIleIleGluArgLeuLeuAlaValProAsnHisLeu 332
Db 183 GACATGGAATCACTACGACATCATCGAGCGCCCTCTGAAGCTGCCGCTCCCATACCTC 242
OY 333 TLeuPheLeuPhePhePheTyrTrpLeuPheHisSerCysLeuAsnAlaValAlaGluLeu 352
Db 243 ATCTGGCTCATCTTCTTCTACTGCGCTCTTCACTCTGCTGATATGCCCTGCTGAGCTC 302
OY 353 MetGlnPheGlyAsnArgGluPheTyrArgAspTrpTrpAsnSerGluSerValThrTyr 372
Db 303 ATGCGATTGAGACCGGAGATTCTACCGGAGCTGGTGAACCTCGAGTCTGTCACCTAC 362
OY 373 PheTrpGlnAsnTrpAsnIleProValHisLysTrpCysIleArgHisPheTyrLysPro 392
Db 363 TTCTGGCAGAACTGGAACTATCCCTGTGCAAGTGGTGATCAGACACTTCTACAAGCCC 422
OY 393 MetLeuArgArgLysSerSerLysTrpMet 402
Db 423 ATGCTTTCGACGGGCGACGACGAATGGATG 452
```

## RESULT 2

```
US-09-867-701-5894
/ Sequence 5894, Application US/09867701
/ Patent No. US20020132237A1
/ GENERAL INFORMATION:
/ APPLICANT: Aglase, Robert A.
/ APPLICANT: Jones, Robert
/ APPLICANT: Harlocker, Susan L.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.497
/ CURRENT APPLICATION NUMBER: US/09/867,701
/ NUMBER OF SEQ ID NOS: 10912
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5894
/ LENGTH: 371
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-867-701-5894
```

## Alignment Scores:

```
Pred. No.: 6,866-60 Length: 371
Score: 612.00 Matches: 123
Percent Similarity: 99.2% Conservative: 0
Best Local Similarity: 99.2% Mismatches: 1
Query Match: 23.6% Indels: 1
DB: 3 Gaps: 0
```

US-10-659-800-6 (1-488) x US-09-867-701-5894 (1-371)

```
OY 186 ValGluSerIleThrProValGlySerLeuLeuAlaLeuMetAlaHisThrIleLeuPhe 205
Db 1 GTTGAGTCTATACCTCCAGTGGGCTCCCTGCTGGGCTGATGAC-CACACCAATCCCTTTC 59
OY 206 LeuLysLeuPheSerTyrArgAspValAsnSerTrpCysArgArgAlaArgAlaLysAla 225
Db 60 CTCAGAGCTTCTCTCAACCGCAGCGCACTCATGTGTCCGACGGGCCCAAGGCT 119
OY 226 AlaSerAlaGlyLysLysAlaSerSerAlaAlaAlaProHisThrValSerTyrProAsp 245
```

```
Db 120 GCCTCTGACGGAGAAAGAGCCAGACAGTCTGCTGCCCGCACACCGTAGACTACCCGGAC 179
OY 246 AsnLeuThrTyrArgAspLeuTyrTrpPheLeuPheAlaProThrLeuCySTyrGluLeu 265
Db 180 AATCTGACTTACCGGAGATCTTACTACTTCTTTCGCCCCACCTTGTCTACAGCTC 239
OY 266 AsnPheProArgSerProArgIleArgLysArgPheLeuLeuArgArgIleLeuGluMet 285
Db 240 AACTTCCCGGCTCTCCCGCATCCGGAAGCGCTTCTGTGTCGACGAGATCTTGAGATG 299
OY 286 LeuPhePheThrGlnLeuGlnValGlyLeuIleGlnGlnTrpMetValProThrIleGln 305
Db 300 CTGTTCTTACCCAGCTCTCCAGTGGGCTGATCCAGCAGTGAATGTCGCCACCATCCAG 359
OY 306 AsnSerMetLys 309
Db 360 AACTCATGAAAG 371
```

## RESULT 3

```
US-09-764-853-314
/ Sequence 314, Application US/09764853
/ Patent No. US20020090672A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: P1206
/ CURRENT APPLICATION NUMBER: US/09/764,853
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - consult PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 939
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 314
/ LENGTH: 433
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (428)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-314
```

```
Alignment Scores:
Pred. No.: 8,266-59 Length: 433
Score: 603.50 Matches: 126
Percent Similarity: 87.2% Conservative: 3
Best Local Similarity: 85.1% Mismatches: 12
Query Match: 23.3% Indels: 7
DB: 3 Gaps: 1
```

US-10-659-800-6 (1-488) x US-09-764-853-314 (1-433)

```
OY 264 GluLeuAsnPheProArgSerProArgIleArgLysArgPheLeuLeuArgIleLeu 283
Db 1 GAGCTCAACTTTCCTCCCGCTCTCCCGCATCCGGAAGCGCTTCTGCTGACGATCTT 60
OY 284 GluMetLeuPhePheThrGlnLeuGlnValGlyLeuIleGlnGlnTrpMetValProThr 303
Db 61 GAGATGCTGTTCTTCAACCACTCCAGGTGGGCTGATCCAGCAGTGAATGATCCCGCAC 120
OY 304 TleGlnAsnSerMetLysProPheLysAspMetAspTyrSerArgIleIleGluArgLeu 323
Db 121 ATCCAGAACTCCATGAAGCCCTTCAGAGCAATGACTACTACGCATCATGAGGCTTC 180
OY 324 LeuLysLeuAlaValProAsnHisLeuIleTrpLeuIlePhePheTyrTrpLeuPheHis 343
Db 181 CTGAGAGCTGGGGGTCCCAATCACTCATCTGGCTCATCTTCTTACTGCTCTTCCAC 240
OY 344 SerCysLeuAsnAlaValAlaGluLeuMetGlnPhe-GlyAspArgGluPheTyrArgAs 363
Db 241 TCTCTCGTAATGCCGTGGCTGAGCTCATGCAAGTTGGAGACCGGAGATTCTACCGGGA 300
OY 363 PTrpTrp-AsnSerGluSerValThrTyrPhe---TrpGlnAsnTrp-AsnIleProVal 381
```

Db 301 CTGGTGGGAATTCGAGTGTTCACCTAATTTCTGGGAGCACTGGGAACATCCCTGTG 360  
Qy 382 -HislystPcysIleArg--HisPheTyrlYsPrometleuArgArgIysSerIys 400  
Db 361 GCACAGTGTGATTCACACATTTTAAAGCCAKGTTCCTGAGGGGAGCAGACAA 420  
Qy 401 TTPMetAlaArg 404  
Db 421 TTGATGCAGCAGG 432

RESULT 4  
US-09-918-995-29256  
; Sequence 29256; Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hybreg, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIORITY APPLICATION NUMBER: US/09/235,076  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 29256  
; LENGTH: 447  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1) ..(447)  
; OTHER INFORMATION: n = A, T, C or G  
US-09-918-995-29256

Alignment Scores:  
Pred. No.: 1,32e-54 Length: 447  
Score: 567.00 Matches: 107  
Percent Similarity: 77.7% Conservative: 1  
Best Local Similarity: 77.0% Mismatches: 0  
Query Match: 21.9% Indels: 31  
DB: Gaps: 2

US-10-659-800-6 (1-488) x US-09-918-995-29256 (1-447)

Qy 298 GINTP-----MetValProThrIleGlnAsnSerMetIysProPheIys----- 312  
Db 31 CGGTGAATTCGCGATGTGTCCCAACATCCATGATGAGCCCTTCAAGGGAGCT 90  
Qy 312 ----- 312  
Db 91 GGCTATGTCTTTCGACGCTGGGGTGGCTGGGGAGTCAACGAGACATGCTAGCTGAA 150  
Qy 313 -----AspMetAspTyrSerArgIleIleGlnArgLeuLeuIysLeuAl 327  
Db 151 GGGCTGTGTTTCGAGGACATGAGTACTACGACATCAACGCGCTCTCTGAAGTGGC 210  
Qy 327 aValProAsnHisLeuIleTyrLeuIlePhePheTyrlYsLeuPheHisSerCysLeuAs 347  
Db 211 GGTCCCAATCACCCTCATCTGCTCATCTTCTTACTAGCTGCTTCCACTCCTGCTGAA 270  
Qy 347 naIaValAlaGlnLeuMetGlnPheGlyAspArgGluPheTyrlYsArgAspTyrTPAsnSe 367  
Db 271 TGCCTGTGCTGAGCTTCATGACATTTGAGAACCGGAGATTCATCCGAGACTGGTGGAACTC 330  
Qy 367 rGluSerValThrTyrlYsPheTyrGlnAsnTyrPAsnIleProValHislystPcysIleAr 387  
Db 331 CGAGCTGTGACCTACTCTTGTGACAGACTGAAATCCCTGTGCAACAAGTGTGATCG 390  
Qy 387 gHisPheTyrlYsPrometleuArgArgIysSerIysTyrMetAlaArgThr 405  
Db 391 ACACTTTCACAGCCCATGCTTCGACGGGGCAGCAGCAAGTGAAGGCCAGACAA 445

RESULT 5  
US-10-741-600-12886  
; Sequence 12886; Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: Cargill, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CLO01499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; PRIOR FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12886  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-12886

Alignment Scores:  
Pred. No.: 4.78e-33 Length: 201  
Score: 374.00 Matches: 65  
Percent Similarity: 98.5% Conservative: 0  
Best Local Similarity: 98.5% Mismatches: 1  
Query Match: 14.4% Indels: 0  
DB: Gaps: 0

US-10-659-800-6 (1-488) x US-10-741-600-12886 (1-201)

Qy 346 LeuAsnAlaValAlaGlnLeuMetGlnPheGlyAspArgGluPheTyrlYsArgAspTyrTP 365  
Db 3 CTGAATGCGGTGGCTGAGCTATGATGAGACCCGGGAACTTACCGGACCTGTGG 62  
Qy 366 AsnSerGluSerValThrTyrlYsPheTyrGlnAsnTyrPAsnIleProValHislystPcys 385  
Db 63 AACTCGAGTCTGTACCTACTTCTGGCAGAACTGGAASATCCCTGTGCACAAGTGTGC 122  
Qy 386 IleArgHisPheTyrlYsPrometleuArgArgIysSerIysTyrMetAlaArgThr 405  
Db 123 ATCAGACCTTTCACAGCCCATGCTTCGACGGGGCAGCAGCAAGTGAAGGCCAGACAA 182  
Qy 406 GlyValPheLeuAlaSer 411  
Db 183 GGGGTGTCTCGCTCG 200

RESULT 6  
US-10-741-600-12898  
; Sequence 12898; Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: Cargill, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CLO01499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; PRIOR FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12898  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-12898

Alignment Scores:  
Pred. No.: 4.78e-33 Length: 201  
Score: 374.00 Matches: 65  
Percent Similarity: 98.5% Conservative: 0  
Best Local Similarity: 98.5% Mismatches: 1  
Query Match: 14.4% Indels: 0  
DB: Gaps: 0

US-10-659-800-6 (1-488) x US-10-741-600-12898 (1-201)

```
OY 346 LeuAenAlaValAlaGluLeuMetGlnPheGlyAspArgLysPheTyrArgAspTrp 365
Db 3 CTGAATGCGCTGGAGTCTCATGACGTTGGAGACGGGAGTTCTACCGGAGCTGGAG 62
OY 366 AenSerGluSerValThrTyrPheTrpGlnAenTrpAsnLeuProValHisLysTrpCys 385
Db 63 AACTCCGAGTCTGTACCTACTTCTGGCAGAACTGGAAATCCCTGTCACAAAGTGTC 122
OY 386 IleArgHisPheTyrLysProMetLeuArgGlySerSerLysTrpMetAlaArgThr 405
Db 123 ATCAGACACTTCTTACAAAGCCCATGCTTCGACGGGACAGCAGCAAGTGATGCCAGACA 182
OY 406 GlyValPheLeuAlaSer 411
Db 183 GGGGTTCTCTGGCTCG 200

RESULT 7
US-09-918-995-1979
; Sequence 1979, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1979
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(471)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1979

Alignment Scores:
Pred. No.: 1.2e-27 Length: 471
Score: 332.00 Matches: 71
Percent Similarity: 53.4% Conservative: 0
Best Local Similarity: 53.4% Mismatches: 2
Query Match: 12.8% Indels: 62
DB: 3 Gaps: 2

US-10-659-800-6 (1-488) x US-09-918-995-1979 (1-471)
OY 315 AapYrSerArgLleIleGluArgLeuLeuLysLeuAla----- 327
Db 58 GACTACTACGCGCATCATCGAGCGCTCTCGAAGCTGGC-GGTGAGTGGCAGACAGTGGCG 116
OY 327 ----- 327
Db 117 CATGCACAGACAGAGGCGACATGTCGGGGGAAAGTTCTAAGACTTGGTCCCA 176
OY 328 -----ValProAsnHisLeuLleTrpLeuLlePhePheTyrTrpLeu 341
Db 177 CCCCACCTCCCTGCGACAGTCCCAATCATCTGCTCATCTTCTTACTGGGCTC 236
OY 342 PheHisSerCysLeuAenAlaValAlaGluLeuMetGlnPheGlyAspArgLysPheTyr 361
Db 237 TTCACCTCCTGCTGAATGCGCTGCTGAGCTCATGCAAGTTTGAGACCGGAGTTCTAC 296
OY 362 ArgAspTrpTrpAsn----- 366
Db 297 CCGGACTGTGTGTG-AGTGTCCCTGGGGGTGTCCTGGGGGCTGGGAGTGGCCATGTGTG 355
OY 367 -----SerGluSerValThrTyrPheTrp 374
```

```
Db 356 CTCTGATCCCCGTGAGTCTCTTGGCCCCCAGAGACTCGAGTCTGTCACTTCTGG 415
OY 375 GlnAenTrpAsnLleProValHisLysTrpCysIleArg 387
Db 416 CAGAACTGGAACTATCCCTGTGTCACAAAGTGATCATCAG 454

RESULT 8
US-09-770-791-192
; Sequence 192, Application US/09770791
; Patent No. US20020062014A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krieker, Maja
; APPLICANT: Slader, Red
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurdan, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2029 (PARA-018PRV)
; CURRENT APPLICATION NUMBER: US/09/770,791
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,480
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(380)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-791-192

Alignment Scores:
Pred. No.: 1.27e-27 Length: 380
Score: 330.50 Matches: 60
Percent Similarity: 69.3% Conservative: 26
Best Local Similarity: 48.8% Mismatches: 34
Query Match: 12.7% Indels: 3
DB: 3 Gaps: 2

US-10-659-800-6 (1-488) x US-09-770-791-192 (1-380)
OY 273 IleArgLysArgPheLeuLeuArgArgLleLeuGluMetLeuPhePheThrGlnLeuGln 292
Db 16 ATACGAAAGGTTGGGTGGTCTGTCATATTCCAAAAGTGCATATTCACCGATTCTATG 75
OY 293 ValGlyLeuLleGlnGlnTrpMetValProThrLleGlnAenSerMetLysProPheLys 312
Db 76 GGATTATTAATAGAACATATATTAATCTTTGTGACAGAACTCAAAAGCATCTTGA 135
OY 313 ---AapMetAapYrSerArgLleIleGluArgLeuLeuLysLeuAlaValProAsnHis 331
Db 136 GCGCATCTTCTATATGCT-----ATTCAAGAGAGTTGAACTTTCAGTTCCAATTTA 189
OY 332 LeuLleTrpLeuLlePhePheTyrTrpLeuPheHisSerCysLeuAenAlaValAlaGlu 351
```

```

Db      190 TAATGCTGGCTTCGACATGTTCATTACTCCTTCTCCACCTTTGGTTAAACAATATTGGCAGAG 249
Qy      352 LdwMetGIInPheGLYAsPaRgIunPhETyrAgaSPtRfRpAsnSergIuseRValThr 371
Db      250 CTTCCTCTCTCTCGGGGATCGTAATTCATCAAAAGATGTGGCATCAAAGAAGTGCGGA 309
Qy      372 TyrPHeTrPGIAAsnTRypAsnIIeRPovAlHslvSTRpCyellleARghIsPheTYrLys 391
Db      310 GATTCACGAGATGATGGNNATRGCTGTTCATTAATGATGATGGTTCGACATATATACCTTC 369
Qy      392 PRoMeLeu 394
Db      370 CGGTGCTTG 378

RESULT 9
US-10-741-600-12888
; Sequence 12888, Application US/10741600
; Publication No. US2005002619A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ.ID NOS: 73997
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO. 12888
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-741-600-12888
```

```

Alignment Scores:
Pred. No.: 5,64e-27 Length: 201
Score: 321.00 Matches: 65
Percent Similarity: 98.5% Conservative: 1
Best Local Similarity: 97.0% Mismatches: 1
Query Match: 12.4% Indels: 0
DB: 8 Gaps: 0

US-10-659-800-6 (1-488) x US-10-741-600-12888 (1-201)

Qy 119 GIhVaIaSeRIePhLeuLyAaPProHSeTTPRoAlaProCySeuVaIIie 138
Db 1 CAGGGGCTTCTCTCTCTCTCTCTGAAGGATCCCTAAGTGGCCCGCCCAAGCTGGTATT 60

Qy 139 AlAlaAenVaIPhelaVaIaAlaIaPheGInaIgluLyArgIeuaIaVAlaIa 158
Db 61 GCGGCAATGCTCTTGTCTGTGGCTGCATCCAGGTTGAGARGGCCTGGCGGGGTGCC 120

Qy 159 LeuThrgIuGInAlGlyLeuLeuLeuHISVaIAlaAenLeuAlaThrlLeuCySPHe 178
Db 121 CTGACGGAGCGAGCGCGGACTGCTGCTGCACGtGGCCAACTGGCACCATTCTGTGTTTC 180

Qy 179 ProAlaAlaVaIleuLeu 185
Db 181 CCAGCGGCTGTGGTCTTACTG 201

RESULT 10
US-10-741-600-12894
; Sequence 12894; Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 12894
;
; LENGTH: 201
;
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-12894

Alignment Scores:
Pred. No.:      5.64e-27      length:      201
Score:          321.00      Matches:      65
Percent Similarity: 98.5%      Conservative: 1
Best Local Similarity: 97.0%      Mismatches:  1
Query Match:    12.4%      Indels:      0
DB:             8          Gaps:         0

US-10-659-800-6 (1-488) x US-10-741-600-12894 (1-201)

```

QY	113	GLNValValSerLeuPheLeuLeuYAspProHisAspSerTPProAlaProCysLeuValIle	138
QY	113	GLNValValSerLeuPheLeuLeuYAspProHisAspSerTPProAlaProCysLeuValIle	138
Db	1	CAAGTGGTTCTCTGTCTCTGAGGAGATCCCTATACCTGAGCCGCCCATGCTGGTATT	60
QY	139	AlaAlaAsnValPheAlaValAlaAlaPheGlnValGluValArgLeuAlaValGlyAla	158
Db	61	GGCGCCAAATGTCCTTTCGTGCTGCATTTCCAGGTGAGAGGGCGCTGGCGGTGGGCC	120
QY	159	LeuThrGluGlnAlaGlyLeuLeuLeuHisValAlaAsnLeuAlaThrIleLeuCysPhe	178
Db	121	CTGACGCGACAGCGCGGAGTCTGCTGTGACGCGGCAACTGGGCACCATCTGTGTTC	180
QY	179	ProAlaAlaValValLeuLeu	185
Db	181	CCAGCGAGCTGTGTTACTG	201

RESULT 11  
US-10-741

```

; Sequence 12900, Application US/10741600
; Publication No. US20050026169a1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12900
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-12900

```

Alignment Scores:	
Pred. No.:	5.64e-27
Score:	321.00
Percent Similarity:	98.5%
Best Local Similarity:	97.0%
Query Match:	12.4%
DB:	8
	Gaps: 0
	Gaps: 0
Length:	201
Matches:	165
Conservative:	1
Mismatches:	1
Indels:	0
Gaps:	0

US-10-659-800-6 (1-488) x US-10-741-600-12900 (1-201)

Oy	119	GlnValValSerLeuPheLeuLysAspProHisSerTyrProAlaProCysLeuValIle	138
		...	
Db	1	CAGTGGTTTCTCTGTTCTCTGAAGATCCCTATAGCTGGCCGCCCATGCTGTTATT	60
Oy	139	AlaAlaAsnValPheAlaValAlaAlaPheGlnValGluLysArgLeuAlaValGlyAla	158
Db	61	GGCGCAATGTCCTTGTCTGTGCTGCATTTCCAGGTTGGAGCGCCCTGGCGGTGGCGCC	120
Oy	159	LeuThrGluGlnAlaGlyLeuLeuLeuHisValAlaAsnLeuAlaThrIleLeuCysPhe	178
		...	
Db	121	CTACCGACACAGCGGGAGTCGTCTGTGACCGGGCAACCTGGCCACCATTCGTGTTTC	180

Qy 179 ProAlaAlaValIleuLeu 185  
|||  
Db 181 CCAGCGGCTGTGCTTACTG 201

## RESULT 12

US-10-741-600-12905  
; Sequence 12905, Application US/10741600  
; Publication No. US20050026169A1

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01499

CURRENT APPLICATION NUMBER: US/10/741,600

NUMBER OF SEQ ID NOS: 73997

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12905

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-10-741-600-12905

## Alignment Scores:

Pred. No.:	5,646-27	Length:	201
Score:	321.00	Matches:	65
Percent Similarity:	98.5%	Conservative:	1
Best Local Similarity:	97.0%	Mismatches:	1
Query Match:	12.4%	Indels:	0
	8	Gaps:	0

US-10-659-800-6 (1-488) x US-10-741-600-12905 (1-201)

Qy 119 GlnValValSerLeuPheLeuLysAspProHisSerTrpProAlaProCysIleuValIle 138  
|||  
Db 1 CAGGGGTTTCTCTCTCTCTGTAAGATCCCTTACTGCGCCGCCCATGCTGTATT 60

Qy 139 AlaAlaSerValPheAlaValAlaAlaPheGlnValGluLysArgLeuAlaValGlyAla 158  
|||  
Db 61 GCGGCAATGTTCTTGTCTGTGCTGCTGATCCAGGTTGAGACGCCCTGCGGTGGTCC 120

Qy 159 LeuThrGlnGlnAlaGlyLeuLeuLeuHisValAlaSerLeuAlaThrIleLeuCysPhe 178  
|||  
Db 121 CTGAGGAGCAGACGGCGGACTGCTGCTGCGACGTGGCCAACTGGCACCATTCTGTGTTTC 180

Qy 179 ProAlaAlaValIleuLeu 185  
|||  
Db 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Qy 181 CCAGCGGCTGTGCTTACTG 201

Pred. No.: 4,496-23 Length: 362  
Score: 290.50 Matches: 55  
Percent Similarity: 73.9% Conservative: 30  
Best Local Similarity: 47.8% Mismatches: 27  
Query Match: 11.2% Indels: 3  
Gaps: 2

US-10-659-800-6 (1-488) x US-10-424-599-115051 (1-362)

Qy 247 LeuThrTrpArgAspLeuTrpPheLeuPheAlaProHisLeuCysTrpGluLeuAsn 266  
|||  
Db 22 GFAACCTTCAAGAGCTTACATATTCTGTGTCCTTACATTTATGTTTACAGCAAGC 81

Qy 267 PheProArgSerProArgIleArgLysArgPheLeuLeuArgArgIleLeuGluMetLeu 286  
|||  
Db 82 TATCTCCACACCTTATATTCGAAAGGTTGGCTGTTCGCCAATGTCACAGCTATA 141

Qy 287 PhePheThrGlnLeuGlnValGlyLeuIleGlnIntPheValProThrIleGlnAsn 306  
|||  
Db 142 ATATTACAGAGATTATGGATTATATATACACATATTAATCCCATTTGACAAAT 201

Qy 307 SerMetLysProPheLys--AspMetAspTrpSerArgIleIleGluArgLeuLeuLys 325  
|||  
Db 202 TCACAGCATCTCTTCAGAGGAACCTTTTACGCC-----ATCGAGAGGTTCTGAG 255

Qy 326 LeuAlaValProAsnIleLeuIleTrpLeuIlePhePheTrpLeuPheHisSerCys 345  
|||  
Db 256 CTTTCTGTTCCAAATTTATATGTTGGCTCGCATGTTCTATTGTTTTCACCTTTGG 315

Qy 346 LeuAsnAlaValAlaGluLeuMetGlnPheGlyAspArgGluPhe 360  
|||  
Db 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Qy 316 TTAATATATATGGCAGAGCTTCTTCGATTGTTGATGCTGTAATTC 360

Tue May 9 11:43:30 2006

RESULT 15

US-10-741-600-12893  
 ; Sequence 12893, Application US/10741600  
 ; Publication No. US20050026169A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C0001499  
 ; CURRENT APPLICATION NUMBER: US/10/741,600  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 73997  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12893  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-600-12893

Alignment Scores:

Pred. No.:	3,41e-23	Length:	201
Score:	288.00	Matches:	54
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.1%	Indels:	0
DB:	8	Gaps:	0

US-10-659-800-6 (1-488) x US-10-741-600-12893 (1-201)

QY	435	MetAlaGlnIleProLeuAlaTrpPheValGlyArgPheGlnGlyAsnTrpGlyAsn	454
DB	3	ATGGCTCAGATCCCACTGGCTGGTTCCGTGGCCCGCTTTTCCAGGGCAACTATGGCAAC	62
QY	455	AlaAlaValTrpLeuSerLeuIleIleGlyGlnProIleAlaValLeuMetTrpValHis	474
DB	63	GCAGCTGTGTGGCTGTGGCTCATCATCGGACAGCCAAATGCGCTCCTCATGTACGTCCAC	122
QY	475	AspTrpTrpValLeuAsnTrpGluAlaProAlaAlaGluAla	488
DB	123	GACTACTACGTGTCTCACTATGAGGCCCGCGGCGAGAGGCC	164

Search completed: May 7, 2006, 03:28:27  
 Job time : 1157 Secs

**THIS PAGE BLANK (uspro)**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 6, 2006, 12:03:36 ; Search time 378 Seconds  
(without alignments)  
3442.262 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MGDGSSRRRTGSRPSHG.....VLMVHYDYVINYEAFAEA 488

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues  
Total number of hits satisfying chosen parameters: 1790828

Minimum DB seq length: 0  
Maximum DB seq length: 500  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-O=abss/ABSSWEB\_spool/US10659800/runat\_05052006.122310.17268/app\_query.fasta\_1  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2n.szlm500.rn1 -MINMATCH=0.1  
-LOOCL=0 -LOOEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=bl0sum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=spct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=500 -HOST=abss02p  
-USER=US10659800@CGN1\_1.512@runat\_05052006.122310.17268 -NCPU=6 -ICPU=3  
-NO MMAP -NCG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCITUS\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	232	8.9	299	US-09-326-203A-13	Sequence 13, Appl
2	197.5	7.6	381	US-09-326-203A-20	Sequence 20, Appl
3	176	6.8	113	US-09-621-976-14231	Sequence 14231, A
4	170	6.6	275	US-09-326-203A-5	Sequence 5, Appl
5	164	6.3	225	US-09-172-711-58	Sequence 58, Appl
6	155	6.0	253	US-09-326-203A-7	Sequence 7, Appl
7	127	4.9	267	US-09-326-203A-4	Sequence 4, Appl
8	123.5	4.8	234	US-09-326-203A-3	Sequence 3, Appl

9	103	4.0	200	US-09-533-559-1684	Sequence 1684, Ap
10	91	3.5	447	US-09-252-991A-292	Sequence 292, App
11	89.5	3.5	325	US-09-326-203A-10	Sequence 10, Appl
12	88.5	3.4	453	US-09-252-991A-11268	Sequence 11268, A
13	88.5	3.4	465	US-09-252-991A-8944	Sequence 8944, Ap
14	88.5	3.4	485	US-09-270-767-10793	Sequence 10793, A
15	88	3.4	257	US-09-326-203A-6	Sequence 6, Appl
16	88	3.4	420	US-09-902-540-5709	Sequence 5709, Ap
17	87.5	3.4	419	US-09-270-767-26258	Sequence 26258, A
18	87	3.4	405	US-09-252-991A-7771	Sequence 7771, Ap
19	87	3.4	462	US-09-902-540-7093	Sequence 7093, Ap
20	86	3.3	354	US-09-252-991A-7022	Sequence 7022, Ap
21	85	3.3	369	US-09-902-540-8920	Sequence 8920, Ap
22	83	3.2	405	US-09-902-540-5868	Sequence 5868, Ap
23	82	3.2	496	US-09-854-133-495	Sequence 495, App
24	81.5	3.1	482	US-09-902-540-6395	Sequence 6395, Ap
25	81	3.1	429	US-09-489-039A-1628	Sequence 1628, Ap
26	80.5	3.1	429	US-09-252-991A-6653	Sequence 6653, Ap
27	80	3.1	312	US-09-902-540-5352	Sequence 5352, Ap
28	80	3.1	420	US-09-252-991A-4348	Sequence 4348, Ap
29	80	3.1	441	US-08-914-375C-51	Sequence 51, Appl
30	80	3.1	486	US-09-252-991A-1276	Sequence 1276, Ap
31	79.5	3.1	418	US-10-131-827-8386	Sequence 8386, Ap
32	79.5	3.1	423	US-09-854-133-713	Sequence 713, App
33	79	3.0	435	US-08-692-922-3	Sequence 3, Appl
34	79	3.0	465	US-09-252-991A-15919	Sequence 15919, A
35	79	3.0	483	US-09-902-540-1856	Sequence 1856, Ap
36	78.5	3.0	464	US-08-718-538-6	Sequence 6, Appl
37	78	3.0	372	US-09-513-999C-17335	Sequence 17335, A
38	78	3.0	336	US-09-902-540-8592	Sequence 8592, Ap
39	78	3.0	479	US-08-757-541-6	Sequence 6, Appl
40	78	3.0	479	US-09-033-275-6	Sequence 6, Appl
41	78	3.0	479	US-09-342-581-6	Sequence 6, Appl
42	77.5	3.0	285	US-09-902-540-9636	Sequence 9636, Ap
43	77.5	3.0	402	US-09-252-991A-10958	Sequence 10958, A
44	77.5	3.0	438	US-09-252-991A-3849	Sequence 3849, Ap
45	77	3.0	269	US-09-313-294A-115	Sequence 115, App

## ALIGNMENTS

RESULT 1  
US-09-326-203A-13  
; Sequence 13, Application US/09326203A  
; Patent No. 6444876  
; GENERAL INFORMATION:  
; APPLICANT: Laseiner, Mike  
; APPLICANT: Ruezinsky, Diane  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; FILE REFERENCE: 17045/00/NO  
; CURRENT APPLICATION NUMBER: US/09/326, 203A  
; CURRENT FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,143  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/108,389  
; PRIOR FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 299  
; TYPE: DNA  
; ORGANISM: murine  
US-09-326-203A-13

Alignment Scores:  
Pred. No.: 2.77e-15  
Score: 232.00  
Percent Similarity: 96.1%  
Best Local Similarity: 86.3%  
Query Match: 8.9%  
DB: 3  
Length: 299  
Matches: 44  
Conservative: 5  
Mismatch: 2  
Indels: 1  
Gaps: 0

```
US-10-659-800-6 (1-488) x US-09-326-203A-13 (1-299)
Qy 434 MetMetAlaGlnIleProLeuAlaTrpPheValGlyArgPheGlnGlyAsnTyrGly 453
Db 3 ATGATGGCTCAGTCCAGCTGCGCTGGATTGGGCCGCAATTTCCAAAGGGAATATGAC 62
Qy 454 AsnAlaValAlaTrpLeuSerLeuIleIleGlyGlnProIleAlaValLeuMetTyrVal 473
Db 63 AATGAGCTGTGTGGTGTGACACTCATTCATTGGGCAACGGGTGGTGTCTCATGTATATGC 121
Qy 474 HisApIYrTyrValLeuAsnTyrGlnAlaPro 484
Db 122 CACGACTACTAGTGTCTCAACTACATGATGCCCA 154

RESULT 2
US-09-326-203A-20
; Sequence 20, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Laesener, Mike
; APPLICANT: Ruzinskiy, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (46)
; OTHER INFORMATION: n at position 46 is unknown
US-09-326-203A-20

Alignment Scores:
Pred. No.: 1,75e-11 Length: 381
Score: 197.50 Matches: 41
Percent Similarity: 61.2% Conservative: 22
Best Local Similarity: 39.8% Mismatches: 31
Query Match: 7.6% Indels: 9
DB: 3 Gaps: 2

US-10-659-800-6 (1-488) x US-09-326-203A-20 (1-381)
Qy 47 PrcAlaProAlaProAsn-----LysAspGlyAspAlaGlyValGly 60
Db 71 CCAGTAGAGCGCTCCCTCATTTGGCACAAAATGTAAATTCGACAGGGAAGATTCAGAAATG 130
Qy 61 SerGlyHsTrpGlnLeuArgCysHisArgLeuGlnAspSerLeuPheSerSerAspSer 80
Db 131 AGAGGACCTTGGGCAAAAGTGTATACACTGCTCAAGATCTATTGTTTCCAGCGATTCCT 190
Qy 81 GlyPheSerAsnTyrArgGlyTyrLeuAsnTrpCysValValMetLeuIleLeuSerAsn 100
Db 191 GGATGGACAAATTTCTCGTGATTTCTTCAATTTGCTATTTGTTGGTACTTTCATAT 250
Qy 101 AlaArgLeuPheLeuGlnAsnLeuIleGlyTyrGlyTyrLeuValAspProIleGlnVal 120
Db 251 GGACCGCGGCACTTGAAATGTGTATCAAAATGTGATTTGATACACACCCCTTCAAGTGG 310
Qy 121 ValSerLeuPheLeuLeuAspProHis-----SerTrpProAlaProCysLeuVal 137
Db 311 ATCTCAACGTTTGTGAGCATCATCTCAATTTGGAGCTGGCCAAATCTTGCTCTCATC 370
Qy 138 IleAlaAla 140
```

```
Db 371 CTATGCTCA 379

RESULT 3
US-09-621-976-14231
; Sequence 14231, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 14231
; LENGTH: 113
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14231

Alignment Scores:
Pred. No.: 4.35e-10 Length: 113
Score: 176.00 Matches: 36
Percent Similarity: 97.3% Conservative: 0
Best Local Similarity: 97.3% Mismatches: 1
Query Match: 6.8% Indels: 1
DB: 3 Gaps: 0

US-10-659-800-6 (1-488) x US-09-621-976-14231 (1-113)
Qy 417 TyrLeuValSerValProLeuArgMetPheArgLeuTrpAlaPheTrnGlyMetMetAla 436
Db 3 TACCTGTGAGCGTCCTCTCGAATGTTCCGCTCTGGGG-TTCACGGGCAATATGCT 61
Qy 437 GlnIleProLeuAlaTrpPheValGlyArgPheGlnGlyAsnTyrGly 453
Db 62 CAGATCCCACTGGCTGTGTGTGTGGCGCGCTTTTCCAGGCACTATGGC 112

RESULT 4
US-09-326-203A-5
; Sequence 5, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Laesener, Mike
; APPLICANT: Ruzinskiy, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (192..(263)
; OTHER INFORMATION: n at positions 192, 202, 204, 211, 222, 234, 238,
; OTHER INFORMATION: 239, 244, 245, 247, 251, 262, and 263 is unknown
US-09-326-203A-5

Alignment Scores:
Pred. No.: 8.01e-09 Length: 275
Score: 170.00 Matches: 34
```

Percent Similarity: 62.7% Conservative: 13  
Best Local Similarity: 45.3% Mismatches: 28  
Query Match: 6.6% Indels: 1  
DB: 3 Gaps: 0

US-10-659-800-6 (1-488) x US-09-326-203A-5 (1-275)

Qy 365 TTPAASerGluSerValThrTyPheTyrGlnAenTyrPasnIleProValHisLysTyr 384  
Db 2 TCGAATGCCAAACCTTGAAGATTTATGAGAGATGSAATATGCTTTCCAAATG 61  
Qy 365 CysIleArgHisPheTyrLysProMetLeuArgGlySerSerLysTyrMetAlaArg 404  
Db 62 ATGATCCGCCACCTATATTCTTCATGTTAAAGCAGGATACCAAGCCGTTGCTT 121  
Qy 405 ThrGlyValPheLeuAlaSerAlaPhePheHisGlyTyrLeuValSerValProLeuArg 424  
Db 122 TTAATTCCTTCCTGCT-CTGCTTATTCATGAGCTGTCGATCGCTGCTTGCCTG 180  
Qy 425 MetPheArgLeuTyrAlaPheThrGlyMetAlaGlnIlePro 439  
Db 181 CATATTCAAGTNGTGCGTTTCNGNGAATTNAGTTTCAGTTCCT 225

RESULT 5

US-09-172-711-58  
; Sequence 58, Application US/09172711  
; Patent No. 6160105  
; GENERAL INFORMATION:  
; APPLICANT: Cunningham, Mary Jane  
; APPLICANT: Zweigler, Gary B.  
; APPLICANT: Panzer, Scott R.  
; APPLICANT: Selhammer, Jeffrey J.  
; TITLE OF INVENTION: MONITORING TOXICOLOGICAL RESPONSES  
; FILE REFERENCE: PA-0011 US  
; CURRENT APPLICATION NUMBER: US/09/172,711  
; CURRENT FILING DATE: 1998-10-13  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PERL Program  
; SEQ ID NO 58  
; LENGTH: 225  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 28, 59, 65, 73, 79, 84, 88  
; OTHER INFORMATION: a or g or c or t, unknown, or other  
; FEATURE:  
; OTHER INFORMATION: 700627089H1  
US-09-172-711-58

Alignment Scores:

Pred. No.: 2,466-08 Length: 225  
Score: 164.00 Matches: 34  
Percent Similarity: 83.0% Conservative: 5  
Best Local Similarity: 72.3% Mismatches: 6  
Query Match: 6.3% Indels: 2  
DB: 3 Gaps: 1

US-10-659-800-6 (1-488) x US-09-172-711-58 (1-225)

Qy 184 LeuLeuValGluSerIleThrProValGlySerLeuLeuAlaLeuMetAlaHisThrIle 203  
Db 85 TTANCGTGAAGTCACTACCTCAGTGGTTCCTGTTGCTTCGCACTACCTCCTC 144  
Qy 204 LeuPheLeuValLeuPheSerTyrArgAlaValAsnSerTyrCys-----ArgArgAla 221  
Db 145 ATCTTCCTCAAGCTTTCTCCCTACCGGAGATGCATCTGTGTGCGCCGAGGAAGGTC 204  
Qy 222 ArgAlaValAlaAlaSerAla 228  
Db 205 AAGGCCAAGCTGTGTCTGCA 225

RESULT 6

US-09-326-203A-7

; Sequence 7, Application US/09326203A  
; Patent No. 6444876  
; GENERAL INFORMATION:  
; APPLICANT: Laesener, Mike  
; APPLICANT: Ruzinskiy, Diane  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; FILE REFERENCE: 17045/00/MO  
; CURRENT APPLICATION NUMBER: US/09/326,203A  
; CURRENT FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,143  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/108,389  
; PRIOR FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 253  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-326-203A-7

Alignment Scores:

Pred. No.: 2,646-07 Length: 253  
Score: 155.00 Matches: 36  
Percent Similarity: 57.6% Conservative: 13  
Best Local Similarity: 42.4% Mismatches: 31  
Query Match: 6.0% Indels: 6  
DB: 3 Gaps: 1

US-10-659-800-6 (1-488) x US-09-326-203A-7 (1-253)

Qy 375 GlnAenTyrPasnIleProValHisLysTyrCysIleArgHisPheTyrLysProMetLeu 394  
Db 1 AGAAATGGAACATGCCCTGTCATMAAGATGTTGCTCATATATATTTCTTGCAAG 60  
Qy 395 ArgArgGlySerSerLysTyrMetAlaArgThrGlyValPheLeuAlaSerAlaPhePhe 414  
Db 61 CGAATGATATATCAAGAGAGTTCCT-TTTATATCGTTCCTGTTCTGCTACTT 119  
Qy 415 HisGlyTyrLeuValSerValProLeuArgMetPheArgLeuTyrAla-PheThrGlyMet 434  
Db 120 CATGATATATGATGCTGCTGCTCCCTGCCACATACAGTTCTGGGCTTTTATGAGAT 179  
Qy 434 tMetAlaGlnIleProLeu-----AlaTyrPheValGlyArgPhePheGlnG 450  
Db 180 CATGCTTCAGATTCCTCCCTCATATATGATGACATCATCACTCAAAAATTAATTCAGTGAC 239  
Qy 450 yAsnTyrGlyAsn 454  
Db 240 AATGCTTGCAAT 252

RESULT 7

US-09-326-203A-4  
; Sequence 4, Application US/09326203A  
; Patent No. 6444876  
; GENERAL INFORMATION:  
; APPLICANT: Laesener, Mike  
; APPLICANT: Ruzinskiy, Diane  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; FILE REFERENCE: 17045/00/MO  
; CURRENT APPLICATION NUMBER: US/09/326,203A  
; CURRENT FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,143  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/108,389  
; PRIOR FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 267

```

; TYPE: DNA
; ORGANISM: Glycine max
US-09-326-203A-4

```

Alignment Scores:	
Pred. No.:	0.000254
Score:	127.00
Percent Similarity:	57.6%
Best Local Similarity:	34.8%
Query Match:	4.9%
DB:	3
	Gaps: 3
	length: 267
	Matches: 32
	Conservative: 21
	Mismatches: 31
	Indels: 9
	Gaps: 3

US-10-659-800-6 (1-488) x US-09-326-203A-4 (1-267)

Qy	192	ValGlySerLeuLeuAlaLeuMetAlaHisThrIleLeuPheLeuLeuPheSerTyr	211
Db	9	GTATCTGGTGCACGTTGATCATTTAACTGCATTTGGGGTTAAATGGTGCATAT	68
Qy	212	ArgAepValAsnSerTyrCysArgArgAlaArgAlaLysAlaLysAlaSerAlaGlyTyrLys	231
Db	69	GCACATACAAAC---TATGATATGAGAGCCTACATGTTTCCAAATGAAAGGAGAA---	122
Qy	232	AlaSerSerAlaAlaAlaProHisThrValSer-----TyrProAspAsnLeuThrTyr	249
Db	123	-----ACATTACCAATACATTGAT-ATGAGATATCCGTAACACTGGACCTTC	169
Qy	250	ArgAepLeuTyrTyrPheLeuPheAlaProThrLeuCysTyrGluLeuAsnPheProArg	265
Db	170	AGGAATTTGGCATCTTCAATGTTGCTCTCACTATTATGCATATGACACAACTATCCCTGC	229
Qy	270	SerProArgIleArgLysArgPheLeuLeuArgArg	281
Db	230	ACACCTTCAGTTTCGAAGGGTTGGGGTTTCGCA	265

RESULT 8  
US-09-32

```

US-09-326-203A-3
Sequence 3, Application US/09326203A
Patent No. 6444876
GENERAL INFORMATION:
APPLICANT: Laesner, Mike
INVENTOR: Ruzinskiy, Diane
TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
Acid Sequences
FILE REFERENCE: 17045/00/NO
CURRENT APPLICATION NUMBER: US/09/326,203A
CURRENT FILING DATE: 1999-06-04
PRIOR APPLICATION NUMBER: 60/088,143
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/108,389
PRIOR FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 234
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc_feature
LOCATION: (24)..(60)
OTHER INFORMATION: n at positions 24, 41, and 60 is unknown
US-09-326-203A-3

```

Alignment Scores:		
Pred. No.:	0.000478	length: 234
Score:	123.50	Matches: 32
Percent Similarity:	67.9%	Conservative: 23
Best Local Similarity:	39.5%	Mismatches: 22
Query Match:	4.8%	Indels: 5
DB:	3	Gaps: 2
US-10-659-800-6 (1-488)	x	US-09-326-203A-3 (1-234)

US-10-659-800-6 (1-488) x US-09-326-203A-3 (1-234)

QY 247 LeuThrTyrArgAspLeuTyrTyrPheLeuPheAlaPro-ThrLeuCysTyrGluLeuAs 266

Db	1	GTAACTTCAGAGAGCTTACGATCATNTTCTCGTGTGGCCCTA	60
QY	266	nPhEProArGserProArGllleArGlyLeArGPhELeuLeuArGa	286
Db	61	CTATCTGCACACCTTATATTGCGAAAGGTTGGCTGTTCCGCA	119
QY	286	nPhEPhEThrGlnLeuGlnValGlyLeuIleGlnInTrpMetVal	306
Db	120	AATATTTCACAGAGCTTATGGATTTATATAGAACATTCATTA	179
QY	306	nSerMetLysPProPheLys---AspMetAspTrpSerArGllle	325
Db	180	TTTACAGATCTCTCTCAAGGGAACCTTTTACCC-----AT	233
QY	325	P 325	
Db	234	G 234	

## RESULT 9

```

US-09-533-559-1684
/ Sequence 1684, Application US/09533559
/ Patent No. 6902887
/ GENERAL INFORMATION:
/ APPLICANT: Randy M. Berka
/ APPLICANT: Michael W. Rey
/ APPLICANT: Jeffrey R. Shuster
/ APPLICANT: Sakari Kauppinen
/ APPLICANT: Ib Groch Clausen
/ APPLICANT: Peter Bjørke Olsen
/ TITLE OF INVENTION: Methods For Monitoring Multiple Gene
/ TITLE OF INVENTION: Expression
/ FILE REFERENCE: 5849, 200-US
/ CURRENT APPLICATION NUMBER: US/09/533,559
/ CURRENT FILING DATE: 2000-03-22
/ EARLIER APPLICATION NUMBER: 09/273,623
/ EARLIER FILING DATE: 1999-03-22
/ NUMBER OF SEQ ID NOS: 7860
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1684
/ LENGTH: 200
/ TYPE: DNA
/ ORGANISM: Fusarium venenatum
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(200)
/ OTHER INFORMATION: n = A,T,C or G
US-09-533-559-1684

```

Alignment Scores:	
Pred. No.:	0.0529
Score:	103.00
Percent Similarity:	52.6%
Best Local Similarity:	34.2%
Query Match:	4.0%
DB:	3
Gaps:	1
Length:	200
Matches:	26
Conservative:	14
Mismatches:	25
Indels:	11

US-10-659-800-6 (1-488) X US-09-533-559-1684 (1-200)

[illegible]

Db 153 TGGCCAAAGCGTGTGGAGAGACTTGGCTCAAGGATTTCATTG 198  
RESULT 10  
US-09-252-991A-292/C  
Sequence 292, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 292  
LENGTH: 447  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-292  
Alignment Scores:  
Pred. No.: 3.62 Length: 447  
Score: 91.00 Matches: 29  
Percent Similarity: 44.0% Conservative: 4  
Best Local Similarity: 38.7% Mismatches: 36  
Query Match: 3.5% Indels: 6  
Gaps: 3  
DB: 3  
US-10-659-800-6 (1-488) x US-09-252-991A-292 (1-447)  
QY 1 MetGlyAspArgGlySerSerArgArgArgGlyTyr-----GlySerArgProSer 17  
Db 429 CTGGGTGATCGCGGCGATCCGACAGCGCCGCGCATCTTGGCGGAGCGCGCTGCC 370  
QY 18 SerHisGlyGlyGlyGlyProHisAlaHisGluGluGluValArgAspAlaHisGly 37  
Db 369 GCT--GGAGTGGCTCGCTGCAACTGTCGCCGACCCGATTTCAGCGACGCCGCTGCCG 313  
QY 38 ProAspValGly-----AlaHisGlyAspAlaProAlaProAlaProHisGlyAspGly 55  
Db 312 GCGCGAAGAGTGGCGCGGCTCAAGCGCGCGCGCGCCGATCTCGGCGAGCGCGCTGCT 253  
QY 56 AspAlaGlyValGlySerGlyHisTyrGluLeuArgCysHisArg 70  
Db 252 GCGCGCGCGCGCGTGCAGTGCAGAGAGTGTGCGACGACGATGATCGC 208  
RESULT 11  
US-09-326-203A-10  
Sequence 10, Application US/09326203A  
Patent No. 6444876  
GENERAL INFORMATION:  
APPLICANT: Laesener, Mike  
APPLICANT: Ruzitsky, Diane  
TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
FILE REFERENCE: 17045/00/WO  
CURRENT APPLICATION NUMBER: US/09/326,203A  
PRIOR FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: 60/088,143  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/108,389  
PRIOR FILING DATE: 1998-11-12  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 325  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:

/ NAME/KEY: misc feature  
/ LOCATION: (6)-(321)  
/ OTHER INFORMATION: n at positions 6, 14, 16, 18, 32, 67, 71, 76, 82,  
/ OTHER INFORMATION: 83, 85, 94, 107, 111, 170, 176, 180, 204, 205,  
/ OTHER INFORMATION: 206, 207, 208, 209, 261, 289, 296, 301, 304, 317,  
/ OTHER INFORMATION: and 321 is unknown  
US-09-326-203A-10  
Alignment Scores:  
Pred. No.: 3.09 Length: 325  
Score: 89.50 Matches: 27  
Percent Similarity: 47.6% Conservative: 12  
Best Local Similarity: 32.9% Mismatches: 33  
Query Match: 3.5% Indels: 10  
Gaps: 4  
DB: 3  
US-10-659-800-6 (1-488) x US-09-326-203A-10 (1-325)  
QY 411 SerAlaPhePheHisGlyTyrLeuValSerValProLeuArgMetPhe-ArgLeuTyrAl 430  
Db 22 TCAGCTGATGCCATGATATGTAATGCGGTGCGCGCCATANTCATTCTN---GG 78  
QY 430 PheThrGlyMetMetAlaGlnIleProLeuAlaTyrPheValGlyArgPhePheGlnI 450  
Db 79 CATNNCGGAGATCATGTTCAGATACCGNTGNA---TTCCTGACACAGATATCTCATGC 135  
QY 450 YAsnTyr-----GlyAspAlaAlaValTyrPheSerLeuIleIleGlyG 465  
Db 136 TACGTTCAAGCATGTAATGTGGGCAACATGTANTTTGGANTCTNACGATATGTGGACA 195  
QY 465 nProIleAlaValLeuMetCysTyrValHisAspTyrTyrValLeuAsnTyrGluAlaProAl 485  
Db 196 GCCGATGNNNNNNATCTATCTACTACATGAC-----GTCTGACACAGCGAGCCAGGC 249  
QY 485 Ala 486  
Db 250 AAGT 253  
RESULT 12  
US-09-252-991A-11268  
Sequence 11268, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 11268  
LENGTH: 453  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-11268  
Alignment Scores:  
Pred. No.: 6.78 Length: 453  
Score: 88.50 Matches: 32  
Percent Similarity: 43.2% Conservative: 3  
Best Local Similarity: 39.5% Mismatches: 23  
Query Match: 3.4% Indels: 23  
Gaps: 5  
DB: 3  
US-10-659-800-6 (1-488) x US-09-252-991A-11268 (1-453)  
QY 4 ArgGlySerSerArgArgArgArgTyrGlySerArgProSerSerHisGlyGlyGly 23  
Db 201 CGTGGCGCGCGCGCGGTGCAATGCCGATACCGGCGCCACCATCATCCTGTGTGGACCT 260

```
QY 24 ProAlaAlaAlaGluGluValAlaArgAspAlaAla----- 36
Db 261 GAAGAGCAGCGCGGATGA-----CCTGCGCTGCGATGACCGACACATCCGCCA 314
QY 37 -----GlyProAspValGly-----AlaAlaGlyAspAla---ProAla 48
Db 315 CTTCGGCAAGGGCCGACCGAAACGAGATACCTGCGCGCGGCGGCAAGCACCTGTCA 374
QY 49 ProAlaProAlaAsp-----GlyAspAlaGlyValGlySercly 62
Db 375 ACTGCTGTGGGCGCAGAAACACGTCGCTGGAGCGGACCTGTGAGTGGAGAGAT 434
QY 63 His 63
Db 435 CAC 437

RESULT 13
US-09-252-991A-8944
; Sequence 8944, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,150
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8944
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8944

Alignment Scores:
Pred. No.: 7.08 Length: 465
Score: 88.50 Matches: 26
Percent Similarity: 44.1% Conservative: 4
Best Local Similarity: 38.2% Mismatches: 29
Query Match: 3.4% Indels: 9
Db: 3 Gaps: 2

US-10-659-800-6 (1-488) x US-09-252-991A-8944 (1-465)
QY 4 ArgGlySerSerArgArgArgThrGlySerArgProSerSerHisGly----- 20
Db 235 AGATCATCCGCAACCGCGCGTGAATCTTCTACACCGACATCACTTTCTGAACGCCT 294
QY 21 -----GlyGlyProAlaAlaAlaGluGluValAlaArgAspAlaAla 35
Db 295 GTCCCTCTCCGCGCGAGAGGAGACCTGCGCACCTCCCGACCTGTCGCGCAGCAGCA 354
QY 36 AlaGlyProAspValGlyAlaAlaGlyAspAlaProAla---ProAlaAsp 54
Db 355 GCCCGCAACAGCTGCGCGCGCGCAACCTCTGCTGCGGCAACCGCGCAACGAG 414
QY 55 GlyAspAlaGlyValGlySercly 62
Db 415 TGGCTGCGCGGCTGCGCACAGC 438

RESULT 14
US-09-270-767-10793/c
; Sequence 10793, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
```

```
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 10793
/ LENGTH: 485
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-10793

Alignment Scores:
Pred. No.: 7.58 Length: 485
Score: 88.50 Matches: 48
Percent Similarity: 41.9% Conservative: 32
Best Local Similarity: 25.1% Mismatches: 68
Query Match: 3.4% Indels: 43
Db: 3 Gaps: 7

US-10-659-800-6 (1-488) x US-09-270-767-10793 (1-485)
QY 43 AlaGlyAspAlaProAlaProAlaProAlaAspGlyAspAlaGlyValGlySercly 62
Db 466 GCTGAGCTCTCCCAATCCGCTTCCCAATGCGAAGTGTGATCATCA----- 416
QY 63 HistrgluLeuArgCysHisArgLeuGluAspSerLeuPheSerSerSerclyPhe 82
Db 416 ----- 416
QY 83 SerAsnTrpArgGlyIleLeuAsnTrpCysValIleLeuIleLeuSerAsnAlaArg 102
Db 415 -----ATTCTTACACCTTGTGTGTTATTTCTGAATGAGCTCAATGCT 371
QY 103 LeuPheLeuGluAsnLeuIleLeuValGlyIleLeuValAspProIleGln---ValVal 121
Db 370 TTGATAGCTTAAATTTGGTATATTTTGGATTTAAATCTGTGAATTCATTCATTCATT 311
QY 122 SerLeuPheLeuAspProHisSerTrpProAlaProCysLeuValIleAlaAsn 141
Db 310 AACCAATT-----TCTGCTCGTCTTAAATTTCAATTAAGCTATCAAT 266
QY 142 ValPheAlaValAlaAlaPheGlnValGluValArgLeuAlaValGlyAlaLeuThr--- 160
Db 265 TTATTTAGAGAGCTTATATACAGATTAATTTTACCTTGTCTAC-----CTTACCTTT 212
QY 161 -----GluGlnAlaGlyLeuLeuLeuHisValAlaAsnLeuAlaThrIleLeuCys 177
Db 211 CAATTAACACAAAGGCTGTTGACTTATTTATTAAGAAATCGTAAAGATCTATTAA 152
QY 178 PheProAlaAlaValAlaLeuLeuValGluSerIleThrProValGlySerLeuLeuAla 197
Db 151 CATGACATGCTGTATTAACATTAAGTAAACCTAATTTGTCAGTCTCAAG 92
QY 198 LeuMetAlaHisThrIleLeuPheLeuIleLeuPheSerTrpArgAspValAsnSerTrp 217
Db 91 TTAATGGCAATGTGCTGGAACCTTTGAAGGACTCTGCTA-----GCATTGAGTTGG 38
QY 218 CysArgArgAlaArgAlaValAlaAlaSerAla 228
Db 37 ---CGCAGATGACAGCGCGTCTGCGCTCA 8

RESULT 15
US-09-326-203A-6
; Sequence 6, Application US/09326203A
; Patent No. 644876
; GENERAL INFORMATION:
; APPLICANT: Laessner, Mike
; APPLICANT: Ruzinskiy, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
```

; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/108,389  
 ; PRIOR FILING DATE: 1998-11-12  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 257  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 US-09-326-203A-6

Alignment Scores:  
 Pred. No.: 3.02 Length: 257  
 Score: 88.00 Matches: 24  
 Percent Similarity: 45.8% Conservative: 14  
 Best Local Similarity: 28.9% Mismatches: 21  
 Query Match: 3.4% Indels: 24  
 DB: 3 Gaps: 5

US-10-659-800-6 (1-488) x US-09-326-203A-6 (1-257)

QY 366 AsnSerGluSerValThrTyrPheTyrGlnAsnTyrPsnIleProValHisIleSTpCys 385  
 Db 43 AATTGTCAACACTGGAGCTTTGGAAAACTGGCATCTTCCTTCAACAAGTGGCTT 102  
 QY 386 IleArgHisPheTyrIleProMetLeuArgGlySerSerIys----- 400  
 Db 103 GTGAGGTATATATACATTCTCTT-----GGGGCATCTAAGAAAAAGCTACTAAT 153  
 QY 401 ---TTPMeValArgThrGlyValPheLeuAlaSerAlaPhePheHisGluTyr----- 417  
 Db 154 GTGTGGGT-----GTTTTCACATTGTGGCAATCTGGCATGATTTAAGAGTG 201  
 QY 418 ---LeuValSerValProLeuArgMetPheArgLeuTyrPalaPheThrGlyMetMetAla 436  
 Db 202 AAGCTTCTTCA-----TGGGCATGGTTGACGTGTTATTC 237  
 QY 437 GlnIlePro 439  
 Db 238 TTCATTCCT 246

Search completed: May 6, 2006, 13:07:20  
 Job time : 381 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 7, 2006, 02:41:49 ; Search time 5784 Seconds  
(without alignments)  
5921.187 Million cell updates/sec

Title: us-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MODRSSRRRTGSRPSSHG.....VLMYHDYVLYNYPAAEA 488

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 32680418

Minimum DB seq length: 0  
Maximum DB seq length: 500

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame.p2n.model -DEV=xlh  
-Q=/abs/ABSSWB\_epool/US10659800/runat\_05052006\_122308\_17225/app\_query.fasta\_1  
-DB=EST-QFMT=fastap -SUFFIX=p2n.szlm500.rst -MINMATCH=0.1 -LOOPTC=0  
-LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.score=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=score -HEAPSIZE=500 -MINLEN=0 -MAXLEN=500  
-HOST=abs802h -USER=US10659800 @CGN\_1.1.5315 @runat\_05052006\_122308\_17225  
-NCPu=6 -ICPU=3 -NO MMAP -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_g981:\*  
10: gb\_g982:\*  
11: gb\_g983:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARY

Result No.	Score	Query Match	Length	DB ID	Description
1	864	33.3	487	10	CZ559383 HMA103 Ba
2	854	32.9	473	3	BM739247 K-EST0008
3	844	32.5	454	3	BM743182 K-EST0016
4	842	32.5	452	1	AA430536 zw21605.r
5	830	32.0	467	2	BG078967 H3034B09
6	798	30.8	485	1	AM446985 88026 MAR
7	762	29.4	492	1	AW702139 uq98d11.x

8	744	28.7	452	1	AM988355 ug07e10.y
9	740	28.5	414	6	CB801052 AMGNNUC.S
10	738	28.5	407	7	CO260334 4132042.B
11	726.5	28.0	459	2	BF733459 RC6-AN006
12	720	27.8	407	6	CB771505 AMGNNUC.T
13	719	27.7	391	6	CB775155 AMGNNUC.T
14	713.5	27.5	434	14	AM391923 QV4-ST023
15	710	27.4	490	3	BM857224 K-EST014
16	707.5	27.3	490	3	BO353862 MR0-HT092
17	705	27.2	457	2	BR844007
18	702	27.1	408	6	CB770458 AMGNNUC.T
19	698	26.9	410	3	BM851541 K-EST0132
20	689	26.6	456	3	BM836400 K-EST0112
21	672	25.9	462	6	CF364199 834132.MA
22	669	25.8	475	2	BE751071 202789.MA
23	666.5	25.7	481	3	BI753832 603027513
24	657	25.3	470	2	BF876883 PM4-ET015
25	656	25.3	479	1	AM446908 87745.MAR
26	655	25.3	470	2	BE647659 UI-M-BH1-
27	653	25.2	410	2	BE868673 601446095
28	649	25.0	375	1	AI182461 uc26h09.r
29	647.5	25.0	484	2	BE308754 601091826
30	646	24.9	422	3	BI340705 366059.MA
31	640	24.7	500	8	N76754 yv45b01.x1
32	635	24.5	380	2	BF884410 CM2-ET019
33	633	24.4	385	7	CN356260 170005325
34	632.5	24.4	416	2	BG314619 OPL_0.125
35	629	24.2	409	6	CB804992 AMGNNUC.N
36	628.5	24.2	406	6	CB807471 AMGNNUC.T
37	627	24.2	443	6	CB452527 707396.MA
38	624.5	24.1	497	3	BP420349 BP420349
39	624	24.1	390	6	CB777920 AMGNNUC.T
40	617	23.8	471	1	AI536527 v08d05.y
41	612	23.6	371	1	AA478000 z034b07.r
42	605.5	23.3	399	1	AA435311 zx47f09.r
43	597	22.0	452	2	BE457297 us49c04.y
44	587	21.6	445	4	AK189743 Mus muscu
45	569	21.9	337	2	BF017684 ux72g03.y

## ALIGNMENTS

RESULT 1  
CZ559383 487 bp mRNA linear GSS 01-JUN-2005  
LOCUS HMA103 BayGenomics Gene Trap Library pGTL1xt Mus musculus cDNA,  
DEFINITION mRNA sequence.  
ACCESSION CZ559383  
VERSION CZ559383.1 GI:66858688  
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 487)  
BayGenomics.  
http://baygenomics.ucsf.edu/  
TITLE Unpublished (2001)  
JOURNAL Contact: BayGenomics  
COMMENT Bay Area Functional Genomics Cohortium (BayGenomics)  
Email: info@baygenomics.ucsf.edu  
Sequence tag generated by 5' RACE of total RNA from gene trap ES  
cell line. ES cell lines harboring insertion mutation of target  
gene are available upon request from BayGenomics. Annotation  
information available from  
http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL.LIN&KEY=HMA103

FEATURES  
source location/Qualifiers  
1..487  
/organism="Mus musculus"  
/mol\_type="mRNA"

/strain="129 oia"  
/db\_xref="taxon:10090"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/clone\_lib="BayGenomics Gene Trap Library pGTLxlf"  
/note="vector: pGTLxlf"

## ORIGIN

## Alignment Scores:

Pred. No.:	3,246-76	Length:	487
Score:	864.00	Matches:	152
Percent Similarity:	96.9%	Conservative:	5
Best Local Similarity:	94.4%	Mismatches:	4
Query Match:	33.3%	Indels:	0
DB:	10	Gaps:	0

US-10-659-800-6 (1-488) x CZ559383 (1-487)

QY 288 PheThrgInLeuGlnValGlyLeuIleGlnGIntPrMeValProThrIleGlnAsnSer 307  
DB 3 TTTACCCAGCTTCAGTGGGGCTGATCCAGAGTGATGCTCCACTACTCCAGAACTCC 62  
QY 308 MetLysProPheLysPheMetAspTyrSerArgIleIleGlnArgLeuLeuLysLeuAla 327  
DB 63 ATGAAGCCCTTCAGAGATGATGACTATTCACCGATCATTCAGACGCTCTTAAAGCTGGCG 122  
QY 328 ValProAsnHisLeuIleTyrPheLeuIlePhePheTyrTyrPheLeuPheHisSerCysLeuAsn 347  
DB 123 GTCCCCAACATCTGATCTGGGCTTACTCTTCTTATTTGGTTTTCACCTCCCTGTCAT 182  
QY 348 AlaValAlaGlyLeuMetGlnPheGlyAspArgGluPheTyrArgAspTyrTrpAsnSer 367  
DB 183 GCTGGGAGAGACTTCTGCAAGTTGGAGACCCGAGTTCTACAGAGATTGGGAGATGCT 242  
QY 368 GluSerValThrTyrPheTyrPheTyrPheLeuIleProValHisLysTyrCysIleArg 387  
DB 243 GAGTCTGTACACTTCTTTGGCAGAACTGGAATATCCCGTGCACAGTGGGCATCACA 302  
QY 388 HisPheTyrLysProMetLeuArgArgGlySerSerLysTyrMetAlaArgThrGlyVal 407  
DB 303 CAGTCTACAGACCTTATGCTTCAGACATGGCAGACGCAAAATGGGTGGCCAGACAGAGAT 362  
QY 408 PheLeuAlaSerAlaPhePheHisGlyTyrLeuValSerValProLeuArgMetPheArg 427  
DB 363 TTTTGGACCTCAGCCTTCTTCCATGATGATCCTAGAGAGCTTCCCTGGAGATGTTCCGC 422  
QY 428 LeuThrAlaPheThrGlyMetMetAlaGlnIleProLeuAlaTyrPheValGlyArgPhe 447  
DB 423 CTCTGGGCATTCACAGCCATGATGGTCCAGTCCCACTGGCCTGGATTTGGGCCGATTC 482  
QY 448 Phe 448  
DB 483 TTC 485

RESULT 2  
BM739247 473 bp mRNA linear EST 01-MAR-2002  
LOCUS BM739247  
DEFINITION K-BST0008669 S2SNU668 Homo sapiens cDNA clone S2SNU668-9-H12 5',  
mRNA sequence.  
ACCESSION BM739247  
VERSION BM739247.1 GI:19060576  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Carcharhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 473)  
Oh,K.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Kim,Y.S., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)

## COMMENT

Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.krrib.re.kr  
Plate: 9 row: H column: 12  
High quality sequence stop: 473.  
Location/Qualifiers

FEATURES  
source

1..473  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S2SNU668-9-H12"  
/sex="M"  
/tissue\_type="Ascites"  
/cell\_type="Epithelial"  
/lab\_host="Top10F"  
/clone\_lib="S2SNU668"  
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;  
Site 2: NotI; The poly (A)+ RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dt-selected  
mRNA by priming with dt-tailed vector. The dt-tailed  
vector was adjusted to have about 60nt. The cDNA vector  
was circularized with E. coli DNA ligase after digestion  
of EcoRI which site is also included in vector. An RNA  
strand converted to a DNA strand by Okayama-Berg method.  
The obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation  
method."

## ORIGIN

## Alignment Scores:

Pred. No.:	3,146-75	Length:	473
Score:	854.00	Matches:	155
Percent Similarity:	98.1%	Conservative:	0
Best Local Similarity:	98.1%	Mismatches:	2
Query Match:	32.9%	Indels:	1
DB:	3	Gaps:	0

US-10-659-800-6 (1-488) x BM739247 (1-473)

QY 245 AspAsnLeuThrTyrArgAspLeuTyrTyrPheLeuPheAlaProThrLeuCysTyrGlu 264  
DB 1 GACAACTGACCTTACCGGATCTTACTTCTTCCGCCCCACCTTGCTTACGAG 60  
QY 265 Leu-AspPheProArgSerProArgIleArgLysArgPheLeuLeuArgAlaLeuGln 284  
DB 61 CTGCAACTTTCGCCGCTCTCCCGCATCCGAAAGGCTTCTGCTGCGACCGATCTTGA 120  
QY 284 UMetLeuPhePheThrGlnLeuGlnValGlyLeuIleGlnGIntPrMeValProThrIle 304  
DB 121 GATGCTGTCTTACACCCAGCTCCAGGTGGGCTGATCCAGAGTGGAGTGGCCACCAT 180  
QY 304 eGlnAsnSerMetLysProPheLysPheMetAspTyrSerArgIleIleGlnArgLeuLeu 324  
DB 181 CCAGAACTCCAGTGAAGCCCTTCAGAGACATGATCTACGACATCATCGAGCGCTCT 240  
QY 324 UlyLeuAlaValProAsnHisLeuIleTyrPheLeuIlePhePheTyrTyrPheLeuPheHis 344  
DB 241 GAACTGGCGGCTCCCAATCACTCATGTGGCTATCTTCTTACTGAGCTTTCAGTC 300  
QY 344 TcysLeuAsnAlaValAlaGlyLeuMetGlnPheGlyAspArgGluPheTyrArgAspTyr 364  
DB 301 CTGCTGAATGGCGGTGGCTGAGTCAATGATGAGACCGGAGTTCATCCGGAGACTG 360  
QY 364 rTrpAsnSerGluSerValThrTyrPheTyrPheTyrPheLeuIleProValHisLysTyr 384  
DB 361 GTGAACTCCAGTCTGTCACTACTTCTGGAGAACTGGAACATCCCTGTGCAAGTGG 420

Qy 384 PCG|lleaRgHisphETyRlySPromETleuARgAGlySerSeryTyr 401  
 |||  
 Db 421 GTCATCAGACACTTCTACAGCCACTGCTGCAGCGGAGACAGCAAGTGG 472

RESULT 3  
 LOCUS BM743182  
 DEFINITION K-EST0016267 S4SNUI Homo sapiens cDNA clone S4SNUI-6-D08 5', mRNA sequence.  
 ACCESSION BM743182  
 VERSION BM743182.1 GI:19064511  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 454)  
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.  
 21C Frontier Korean EST Project 2001  
 TITLE Unpublished (2002)  
 JOURNAL Contact: Kim YS  
 COMMENT Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52, Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.krrib.re.kr  
 Plate: 6 row: D column: 08  
 High quality sequence stop: 454.

# FEATURES

source  
 1..454  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S4SNUI-6-D08"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Lymphoblast-like"  
 /cell\_line="SNU-1"  
 /lab\_host="Top10P"  
 /clone\_lib="S4SNUI"  
 /note="Organ: Stomach; Vector: pTZ189P1; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then deapped with tabacco acid pyrophosphatase (TAP). The deapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoR I which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2.98e-74 Length: 454  
 Score: 844.00 Matches: 150  
 Percent Similarity: 100.0% Conservatave: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 32.5% Indels: 0  
 DB: 3 Gaps: 0

US-10-659-800-6 (1-488) x BM743182 (1-454)

Qy 312 LysAepMeCAspTyRSerArgllellelgUARgleuLeuYaleuAlaValProAsnHis 331  
 |||  
 Db 3 AAGACATGAGATCTACTACAGCATCATCGAGCCCTCTGAACCTGGCCGTCCCATCAC 62

Qy 332 LeuIleTPLeuIlePhpheThyRTrpLeuPheHisSerCysLeuAsnAlaValAlaGlu 351  
 |||  
 Db 63 CTCATCTGGCTCATCTCTCTTCTACTGCGCTTCCACTCCTGCTGATCCCTGGCTGAG 122

Qy 352 LeuMetGlnPheGlyAspArgGluPheThyTrpArgAspTrpAsnSerGluSerValTyr 371  
 |||  
 Db 123 CTCATGAGATTGGAGACCGGAGATTCTACCGGAGCTGGTAGACTCGAGTCTGTACCC 182

Qy 372 TyrPheTrpGlnAspTrpAsnIleProValHisLysTrpCysIleArgHisphETyRlys 391  
 |||  
 Db 183 TACTCTGGCAGAACTGGAACATCCCTGTGCAAGAGTGTGATCAGACACTTCTACAG 242

Qy 392 ProMetLeuArgArgGlySerSerLysTrpMetAlaArgThyGlyValPheLeuAlaSer 411  
 |||  
 Db 243 CCCATGCTTCGACGGGGAGAGCAAGTAGATGAGCAGGACAGGGAGTTCCTGGACCTCG 302

Qy 412 AlaPhePheHisGluTyrLeuValSerValProluArgMetPheArgLeuTrpAlaPhe 431  
 |||  
 Db 303 GCTTCTTCCACGAGTACCTGTGTGAGCGCTCTGCGAATGTTCCGCTCTGGGCGTTC 362

Qy 432 ThrGlyMetMetAlaGlnIleProLeuAlaTrpPheValGlyArgPhePheGlnGlyAsn 451  
 |||  
 Db 363 ACGGCATGATGATGGCTCAATCCCACTGGCTGTGTCGTGGCGCTTTTTCAGGGGCAAC 422

Qy 452 TyrGlyAsnAlaAlaValAlaTrpLeuSerLeu 461  
 |||  
 Db 423 TATGGCAACGACGAGTGTGGCTGTGCTC 452

RESULT 4  
 AA430536 452 bp mRNA linear EST 09-NOV-1997  
 LOCUS ZW21A05.11 Soares ovary tumor NbHOT Homo sapiens cDNA clone  
 IMAGE:769904 5' similar to TR:G1066810 G1066810 ACYL-COENZYME A:  
 CHOLESTEROL ACYLTRANSFERASE. ;, mRNA sequence.  
 ACCESSION AA430536  
 VERSION AA430536.1 GI:2111170  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 452)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheinberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
 WashU-NCI human EST Project  
 Unpublished (1997)  
 TITLE Contact: Wilson RK  
 JOURNAL  
 COMMENT Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1810  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LML ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert length: 1033 Std Error: 0.00  
 Seq primer: -28m13 rev2 ET from Amerham.

## FEATURES

source  
 1..452  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:5979779"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:769904"  
 /sex="Female"  
 /tissue\_type="Ovarian tumor"  
 /lab\_host="DH10B (ampicillin resistant)"



Db 182 GACCGGAGTTCTACAGAGATTGGTGAATGCTGACTCTGCTACCTACTTTTGGCAGAAC 241

Qy 377 TTPaenIIeProValHisIstYrPcysIIeArgHisPheTyrIlyProMetIeuAArg 396

Db 242 TGGAAATATCCCGTGCACAGATGTGATCAGACACTTCTACAGGCTATGCTCAACACAT 301

Qy 397 GlycerSerIleTyrMetAlaArgThrGlyValPheIeuAlaSerAlaPhePheHisGly 416

Db 302 GGCAGAGCAATATGGGTGGCCAGACAGAGATATTTTACACCTCACCCTTTCCATGAG 361

Qy 417 TTTleuValSerValProIeuAArgMetPheArgIleuTTPAlaPheThrGlyMetVal 436

Db 362 TACCTAGTAGAGGCTTCCCTCCGCGATGTTCCGCTCTGGGCATTCACAGCCATGATGGCT 421

Qy 437 GlnIleProIeuAlaTTPheValGlyArgPhePheGlnGlyAsn 451

Db 422 CAGGTCCCACTGACCTGATGTGGCCGCAATTCCTTCCAGGCAAC 466

RESULT 6

AM446985 485 bp mRNA linear EST 25-APR-2001

LOCUS 88026 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AM446985

VERSION AM446985.1 GI:6988772

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecota; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 485)

AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Caste,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Percec,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

PubMed 11282978

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 20 and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG

Plate: 54 row: B column: 13

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1. 485

/organism="Bos taurus"

/mol\_type="mRNA"

/db\_xref="taxon:9913"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_1lb="MARC 1BOV"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

ORIGIN

Alignment Scores:

Pred. No.: 1,38e-69 length: 485

Score: 798.00 Matches: 149

Percent Similarity: 85.5% Conservative: 4

Best Local Similarity: 83.2% Mismatches: 8

Query Match: 30.8% Indels: 18

DB: 1 Gaps: 2

US-10-659-800-6 (1-488) x AM446985 (1-485)

Qy 202 ThrIleIeuPheIeuIleuPheSerIYrArgAspValAsnSerTTPcysArgAla 221

Db 3 ACCATCTCTTCTCCAAAGCTGTCTCTACCCGGAGCTCAACCTCTGTCGCCAGAGCGC 62

Qy 222 ArgAlaIyAlaIaIaSerIaGlyIyIyAlaIaSerSerAlaIaIaIaProHisThrVal 241

Db 63 AGS-----GCTGGGCCAAGGCCAAGCGCTTGGCA-----95

Qy 242 SerTyrProAspAsnIeuThrTyrArgAspIeuTyrTyrPheIeuPheAlaProThrIleu 261

Db 96 -----GATCTCTACTTCTCTCTTCTTCCGCCACCCG 128

Qy 262 CysTyrGluIleuAsnIeuPheProIArgSerProArgIleArgIyAlaArgPheIeuIeuArg 281

Db 129 TGTACAGAGCTCAACTCCCGCTCCCGCTCCCGCATCCGAAGCGCTTCTGCTGCGCGCA 188

Qy 282 IleIeuGluMetIeuPhePheThrIleuGluIleuGlyIleuIleGlnIleuIleuMetVal 301

Db 189 CTCTGAGATGCTGTCTCTCAACCGCTCCAGGTGGGCTATCCAGAGTGGATGTC 248

Qy 302 ProThrIleGlnAsnSerMetIySPropheIyAspMetAspTyrSerArgIleIleGlu 321

Db 249 CCGGCATCCAAATCCCATGAAGCCCTTCAAGACATGATACCTCCGCATCGAGAG 308

Qy 322 ArgIleuIleuIleuAlaValProAsnHisIleuIleTyrIleuIlePhePheTyrTTPleu 341

Db 309 CCCTCTCTGAACCTGGCGGCTCCCAACCACTCATCTGCTCATCTTCTTACTGCGCTC 368

Qy 342 PheHisSerCysIleuAsnAlaValAlaGluMetIeuPheGlyAspArgGluPheTyr 361

Db 369 TTCCACTCTGTGCTGAAGCCCGTGTGCTCATGACATTTGGAAACCGCAGATTCTAC 428

Qy 362 ArgAspTyrTTPAsnSerGluSerValThrTyrPheTyrGlnAsnTTPAsnIlePro 380

Db 429 CCGGACTGTGAACTCCGAGTCATCACTTCTTGGCABAATCGAACAATCCCT 485

RESULT 7

AM702139 492 bp mRNA linear EST 18-APR-2000

LOCUS UW98d11.x1 NCI CGAP Mam10 Mus musculus cDNA clone IMAGE:2939157 3'

DEFINITION similar to TR:Q92ZA7 Q92ZA7 DIACYLGlycerol ACYLTRANSFERASE. ;, mRNA sequence.

ACCESSION AM702139

VERSION AM702139.1 GI:7586288

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 492)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Lohar Hennighausen Ph.D., Jeffrey Green M.D., Gilbert Smith, Ph.D., William Muller, M.D., Jeffrey Green M.D., cDNA Library Preparation: M. Benito Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (IIML)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/IIML at:

image.llnl.gov/image/html/resources.shcml

MGI:1051577

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Glibco  
High quality sequence stop: 448.  
Location/Qualifiers

FEATURES  
source

1..492  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:2939157"  
/tissue\_type="pooled mammary gland tumors"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NCI\_CGAP\_Mam10"  
/note="Organ: mammary gland; Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker; Site 1: NotI;  
Site 2: EcoRI; 1st strand cDNA was prepared from mRNA  
obtained from pooled mammary gland tumors with a Not I -  
oligo(dT) primer (5'  
TGTTACCAATCTGAAGTGGAGCGCGCAGTCTTTTCTTTTCTTTT 3').  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.: 5,886-66 Length: 492  
Score: 762.00 Matches: 131  
Percent Similarity: 97.2% Conservative: 8  
Best Local Similarity: 91.6% Mismatches: 4  
Query Match: 29.4% Indels: 0  
DB: 1 Gaps: 0

US-10-659-800-6 (1-488) x AM702139 (1-492)

Oy 342 PheHISerCYeLeuAaAlaValAlaGluLeuMetGlnPheGlyAspArgGluPheTyr 361  
Db 6 TTCCACTCTCTGCTCAATGCTGTGGCAGAGCTTGTGAGTTGGAGACCGCAGTTCTAC 65  
Oy 362 ArgAspTPPTTPAAsenSerGluSerValThrTyrPheThrGlnAsnTrpAsnIleProVal 381  
Db 66 AGAGATTGGTGAATGCTGAGTCTGTCACTTCTTTGGCAACCTGGAAATATCCCGTG 125  
Oy 382 HisIstTPCYaIleArgHisPheTyrLysProMetLeuArgArgGlySerSerLysTrp 401  
Db 126 CACAAGTGTGATCAGACACTTCTACAGCCTTAGCTAGACAGGACGACGACAAATGG 185  
Oy 402 MetAlaArgThrGlyValPheLeuAlaSerAlaPhePheHisGluTyrLeuValSerVal 421  
Db 186 GTGGCCAGGACGAGAGTATTTTGTGACCTCAGCCTTCTCCATGAGTACCTAGAGCGTT 245  
Oy 422 ProLeuArgMetPheArgLeuTrpAlaPheThrGlyMetMetAlaGlnIleProLeuAla 441  
Db 246 CCCCTGGCGAGTATTCGCTCTGGGCAATTCACAGCCATGATGCTCAGGTCCTCCACTGGCC 305  
Oy 442 TrpPheValGlyArgPhePheGlnGlyAsnTyrGlyAsnAlaAlaValTrpLeuSerLeu 461  
Db 306 TGGATTGTGGCCGCAATTTCTTCCAGGGAATGGAATGCAAGCTGTGGGTGACACATC 365  
Oy 462 IleIleGlyGlnProIleAlaValLeuMetTyrValHisAspTyrTyrValLeuAsnTyr 481  
Db 366 ATCATTTGGCAACCGGTGCTGTGCTCATGTATGTCCAGACTACTACGTGTCAACTAC 425  
Oy 482 GlnAlaPro 484  
Db 426 GATGCCCA 434

RESULT 8  
LOCUS AM988355 452 bp mRNA linear EST 02-JUN-2000  
DEFINITION ugm7e10.y1 Soares\_mammary\_gland\_NMLNG Mus musculus cDNA clone  
IMAGE:1530954 5' similar to TR:0922A7 0922A7 DIACYLGLYCEROL  
ACYLTRANSFERASE. /, mRNA sequence.  
ACCESSION AM988355

VERSION AM988355.1 GI:8183481  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL  
COMMENT Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
MGI:947054

Seq primer: -40RP from Glibco.

## FEATURES

## source

1..452  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1530954"  
/sex="Female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"  
/clone\_1lb="Soares mammary gland\_NMLNG"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified pT73 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,336-64 Length: 452  
Score: 744.00 Matches: 128  
Percent Similarity: 97.1% Conservative: 8  
Best Local Similarity: 91.4% Mismatches: 4  
Query Match: 28.7% Indels: 0  
DB: 1 Gaps: 0

US-10-659-800-6 (1-488) x AM988355 (1-452)

Oy 345 CysLeuAaAlaValAlaGluLeuMetGlnPheGlyAspArgGluPheTyrArgAspTrp 364  
Db 3 TGTCTCATGCTGTGGCAGAGCTTCTGTGAGACCGCGAGTTCTACAGAAATTGG 62  
Oy 365 TrpAAsenSerGluSerValThrTyrPheThrGlnAsnTrpAsnIleProValHisLysTrp 384  
Db 63 TGGAAATGCTGAATCTGTCACTTCTTTGGCAACTGGAATATCCCGTGCACAAATGG 122  
Oy 385 CysIleArgHisPheTyrLysProMetLeuArgArgGlySerSerLysTrpMetAlaArg 404  
Db 123 TGCATCACAACACTTCTTACAACCTATGCTCAGACATGCAACAGCAAAATGGGTGCCAGG 182  
Oy 405 ThrGlyValPheLeuAlaSerAlaPhePheHisGluTyrLeuValSerValProLeuArg 424  
Db 183 ACAGAGATATTTTGAACCTCAGCCTTCTCCATGATGATGAGTTCAGGCTTCCCTGGCG 242  
Oy 425 MetPheArgLeuTrpAlaPheThrGlyMetMetAlaGlnIleProLeuAlaTrpPheVal 444  
Db 243 ATGTTCCGCTCTGGGCAATTCACAGCCATGATGCTCAGGTCCTGACCTGGATTGTG 302  
Oy 445 GlyArgPhePheGlnGlyAsnTyrGlyAsnAlaAlaValTrpLeuSerLeuIleIleGly 464  
Db 303 GGCAGATTTCTTCCAAAGGAACATATGCAATGCAAGCTGTGTGGGTGACACTCATATTGGG 362  
Oy 465 GlnProIleAlaValLeuMetTyrValHisAspTyrTyrValLeuAsnTyrGlnAlaPro 484

Db 363 CAACCGTGGCTGCTCATGATGTCACGACTACTAGTGTCTCACTAGCAATGCCCA 422

RESULT 9  
LOCUS CB801052 414 bp mRNA linear EST 16-MAY-2003  
DEFINITION AMGNNUC:SRBP2-00118-D4-A srpb2 (10220) Rattus norvegicus cDNA clone  
CB801052  
ACCESSION CB801052  
VERSION CB801052.1 GI:29913930  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciuromorphi; Muridae; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 414)  
AUTHORS Amgen EST Program.  
TITLE Amgen Rat EST Program  
JOURNAL Unpublished (2003)  
COMMENT Contact: Dan Fitzpatrick  
Amgen, Inc  
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00118 row: d column: 4.  
Location/Qualifiers  
1..414  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="srpb2-00118-d4"  
/tissue\_type="prostate tissue"  
/clone\_lib="srpb2 (10220)"  
/note="Vector: pSPORT1. Site 1: SalI; Site 2: NotI; rat  
prostate normalized double selected poly(A+) mRNA size  
fraction > 1 Kb"

ORIGIN  
source  
Alignment Scores:  
Pred. No.: 7 37e-64 Length: 414  
Score: 740.00 Matches: 129  
Percent Similarity: 98.5% Conservative: 6  
Best Local Similarity: 94.2% Mismatches: 2  
Query Match: 28.5% Indels: 0  
Gaps: 0

US-10-659-800-6 (1-488) x CB801052 (1-414)

QY 303 Thrllglnasnsermetlyspropheylsaspmetaspyrserarglleilegiuarg 322  
Db 3 ACTATCCAGAACTCCATGAGCCCTTCAGAGACATGAGACTATTCAAGATCATTTAGCGCT 62  
QY 323 LeuLeuLyLeuAlaValaProAnhiSleuIlleTrrPheuIllePhePheTrrPhePhe 342  
Db 63 CTTTAAAGCTGGCGCTCCCAACCATCGATATGCTCATCTTCTTATTTGGCTTTTC 122  
QY 343 HisserCysleuasnAlaValaAlaGluLeuMetGlnPheGlyAspArgGluPheTrrArg 362  
Db 123 CACTATGCTCTAAAGCTGTGGAGAGCTGCGAGTTGGAGACCGCAGATTCTACAG 182  
QY 363 AspTrrPasnserGluSerValThrTrrPheTrrGlnAsnTrrPasnIlleProValHis 382  
Db 183 GACTGTGGAAATCGAGCTGTCTACCTTCTTGGCAGAACTGGAAATCCCGGCGAC 242  
QY 383 LybTrrCysIleArgHisPheTrrLyseProMetLeuArgArgGlySerSerLyseTrrMet 402  
Db 243 AAGGTGTGATCAGACACTTTCACAAAGCTATGCTCAGACTGGGCGAGCAAAATGGATG 302  
QY 403 AlaacGtrngLyvalPheleuAlaserAlaPhePheHisGluTrrLeuValaserValPro 422  
Db 303 GCCAGAGCTGGGGCTTTTGGCGTACGCTTCTTCATAGACTACTAGACATTTCC 362  
QY 423 LeuArgMetPheArgLeuTrrAlaPheThrngLyMetMetAlaGlnIlePro 439

Db 363 CTGAGGATGTTCGCTCTGGGCAATTCACAGCATGATGCTCAGGTCCCA 413

RESULT 10  
LOCUS CO260334 407 bp mRNA linear EST 23-JUN-2004  
DEFINITION 4132042 BARC 8BOV Bos taurus cDNA clone 8BOV\_57F07 5', mRNA  
sequence.  
ACCESSION CO260334  
VERSION CO260334.1 GI:49145172  
KEYWORDS EST.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 407)  
AUTHORS Baumann,R.G., Baldwin,R.L., Sonstegard,T.S., Van Tassel,C.P. and  
Matukumalli,L.K.  
TITLE Construction and Analysis of a cDNA Library Generated From  
Intestinal Muscle and Epithelial Tissues of Holstein Cattle  
JOURNAL Unpublished (2004)  
COMMENT Contact: Richard G. Baumann  
Bovine Functional Genomics Lab  
ANRI  
BLDG 162: BARC-EAST, Beltsville, MD 20705, USA  
Tel: 3015048604  
Fax: 3015048744  
Email: rbaumann@anri.barc.usda.gov  
Single pass sequencing. Bases called and trimmed with phred  
0.000925 using options -trim\_alt '- -trim\_fast. Vector identified  
by cross match using options -mismatch 12 -mnscore 18  
Plate: 57 row: F column: 07  
Seq primer: CCTATTAGTGCACCTATAGAAC  
High quality sequence stop: 407.  
Location/Qualifiers  
1..407  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/strain="Holstein"  
/db\_xref="taxon:9913"  
/clone="8BOV\_57F07"  
/sex="Female"  
/tissue\_type="Epithelial, Muscle"  
/dev\_stage="Lactating, Neonatal"  
/lab\_host="PH10B Tona"  
/clone\_lib="BARC 8BOV"  
/note="Organ: Intestine; Vector: pCMVSPORT6.1; Site 1:  
NotI; Site 2: EcoRI; Normalized cow cDNA intestinal  
library in pCMVSPORT6.1, constructed from equimolar mRNA  
pools derived from 5 sources; 4 lactating intestinal, 1  
neonatal intestinal 4/5 lactating, proximal duodenum,  
dujunum, distal ileum, colon, 1/5 neonatal, proximal  
duodenum, dujunum, distal ileum"

ORIGIN  
source  
Alignment Scores:  
Pred. No.: 1 14e-63 Length: 407  
Score: 738.00 Matches: 128  
Percent Similarity: 97.8% Conservative: 3  
Best Local Similarity: 95.5% Mismatches: 3  
Query Match: 28.5% Indels: 0  
Gaps: 0

US-10-659-800-6 (1-488) x CO260334 (1-407)

QY 309 LysProphelysaspmetaspyrserarglleilegiuargleuLyLeuAlaVal 328  
Db 2 AAGCCCTTCAGAGCAATGACTATTCGCCGACATCGAGCGGCTCTGTAAGCTGGCGGCT 61  
QY 329 ProAnhiSleuIlleTrrPheuIllePhePheTrrPhePheHisSerCysleuAlaVal 348  
Db 62 CCACACCACTCATGCTGCTCATCTTCTTACTGCTCTTCCACCCCTGCTGAGACGCC 121

QY 349 ValAlaGluLeuMetGlnPheGlyAspArgGluPheTyrArgAspTyrTrpAsnSerGlu 368  
Db 122 GTGGCTGAGCTCATCAGTTGGAGACCGGAGTTCTACCGGAGCTGGGAACTCCGAG 181  
QY 369 SerValThrTyrPheTrpGlnAsnTrpAsnIleProValIleGlyTrpCysIleArgHis 388  
Db 182 TCATCACCCTACTTCTGGCAGAACTGGAACTCCCTGTTCCAAAGTGGATCAGAC 241  
QY 389 PheTyrIleProMetIleuArgArgIleSerSerIleTyrMetAlaArgThrIleValPhe 408  
Db 242 TTCTCAAGACCATCTCCGCGGGGAGCAAGTGGGAGCCGAGGAGCGCAGTGT 301  
QY 409 LeuAlaSerAlaPhePheIleGluTyrLeuValSerValProLeuArgMetPheArgLeu 428  
Db 302 CTGGCTCCGCTCTTCTCCACGAGTACTGTGAGAGTCCCTCTGGCGCATGTTCCGCTTC 361  
QY 429 TrpAlaPheThrGlyMetMetAlaGlnIleProLeuAlaTyr 442  
Db 362 TGGGCTTCAACCGCATGATGGCGCAGATCCCGCTGGCTGG 403  
RESULT 11  
BF733499/c 459 bp mRNA linear EST 09-JAN-2001  
LOCUS RC6-AN0063-081000-011-B03 AN0063 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF733499  
VERSION BF733499.1 GI:12058735  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Euarchoptogilres; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 459)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,  
Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,  
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,  
Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V.,  
O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and  
Simpson, A. J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
PUBMED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FADESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC6&t2=RC6-AN0063-  
081000-011-B03&t3=2000-10-08&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 26  
High quality sequence stop: 458.  
Location/Qualifiers  
1..459  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="AN0063"  
/note="Organ: amnion normal; Vector: puc18; Site 1: Smal;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN  
Alignment Scores:  
Pred. No.: 1,96e-62 Length: 459  
Score: 726.50 Matches: 147  
Percent Similarity: 98.7% Conservative: 1  
Best Local Similarity: 98.0% Mismatches: 1  
Query Match: 28.0% Indels: 1  
DB: 2 Gaps: 1  
US-10-659-800-6 (1-488) x BF733499 (1-459)  
QY 103 LeuPheLeuGluAsnLeuIleGlyTyrGlyIleLeuValaProIleGlnValaSer 122  
Db 457 TTATTCTCGAGAACCTCATCAGATGGCATCTCGGTGGAGACCCATCCAGTGGTTCT 398  
QY 123 LeuPheLeuValaProHisSerTyrProAlaProCysLeuValIleAlaAlaSerVal 142  
Db 397 CTGTTCTGAAGAGTCTCTATAGCTGGCCGCCCATGCTGTATTGGCGCCATGTC 338  
QY 143 PheAlaValAlaAlaPheGlnValaGluValaArgLeuAlaValaGlyAlaLeuThrGluGln 162  
Db 337 TTGCTGTGGCTGATCTCAGGTTGAGAAAGCGCTGGCGGTGGTCCCTTACGAGACAG 278  
QY 163 AlaGlyLeuLeuLeuHisValaAlaAsnLeuAlaThrIleLeuGlyPheProAlaAlaVal 182  
Db 277 GCGGGA---CTGCTGCACGTGGCCAACTGGCCACCATTTGTGTGTTCCAGCGGCTGTG 221  
QY 183 ValLeuLeuValaGluSerIleThrProValaGlySerLeuLeuAlaLeuMetAlaHisThr 202  
Db 220 GTCTTACGTGGTGAATCATGCTTCAGTGGGCTCTCCGTGGCCCTATGGCGCACACC 161  
QY 203 IleLeuPheLeuValaLeuPheSerTyrArgAspValaAsnSerTyrCysArgArgAlaArg 222  
Db 160 ATCTCTTCTCAGACCTCTTCTCTCAACCGCAGCACTCATGTGCGGAGGCGCAGG 101  
QY 223 AlaValaAlaSerAlaGlyValaValaAsnSerAlaAlaAlaProHisThrValSer 242  
Db 100 GCCAAGGCTGCTCTGCAGGGAAGAAAGCCAGCATGCTGTGCCCCCAGACCGTGAGC 41  
QY 243 TyrProAspAsnLeuThrTyrArgAspLeu 252  
Db 40 TACCGGACAACTACCTACCGCATCTA 11  
RESULT 12  
CB771505 407 bp mRNA linear EST 16-MAY-2003  
LOCUS AMGNNUC:TXXP1-00014-D10-A txp1 (10556) Rattus norvegicus cDNA  
DEFINITION clone txp1-00014-d10 5', mRNA sequence.  
ACCESSION CB771505  
VERSION CB771505.1 GI:29859896  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Euarchoptogilres; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 407)  
AUTHORS Amgen EST Program.  
JOURNAL Amgen Rat EST Program  
COMMENT Unpublished (2003)  
CONTACT: Dan Fitzpatrick  
Amgen, Inc  
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00014 row: d column: 10.  
Location/Qualifiers  
1..407  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="txp1-00014-d10"  
/tissue\_type="prostate"



ORIGIN /note="Vector: pYVA-41L; Site\_1: HindIII; Site\_2: NotI;  
rat prostate"

Alignment Scores:  
Pred. No.: 7,366-62 Length: 407  
Score: 720.00 Matches: 10  
Percent Similarity: 98.5% Conservative: 4  
Best Local Similarity: 95.6% Mismatches: 1  
Query Match: 27.8% Indels: 1  
DB: 6 Gaps: 0

US-10-659-800-6 (1-488) x CB771505 (1-407)

QY 265 LeuAnpPheProArXSeProArXglLeArXglYsArXpHeuLeuArXglLeuGlu 284  
DB 1 CTCACCTTCTCCGATCCCGCAATACGAAAGCGCTTCTGACGGGGGCTTGTGAG 60  
QY 285 MetLeuPhePheThrGlnLeuGlnValGlyLeuIleGlnGlnTrpMetValProThrIle 304  
DB 61 ATGCTCTTTTACCCAGCTTCAAGTGGGGCTGATCCAGCATGATGATGCTCTACTT 120  
QY 304 eGlnAsnSerMetLysProPheLysAspMetLysPheTyrSerArgIleIleGlnArgLeuLeu 324  
DB 121 CCAAGACTCCATGAGACCCCTTCAAGAGACATGAGATTCACGATCATTTGACGCTCTT 180  
QY 324 ULyLeuAlaValProAsnHisLeuIleTrpLeuIlePhePheTyrTrpLeuPheHisSe 344  
DB 181 AAGCTGGGGGCTCCCAACCATGTGATAGGCTCATCTTCTTATGGCTTTTCCACATC 240  
QY 344 CyLeuAlaValAlaGluLeuMetGlnPheGlnYsAspArgGluPheTyrArgAspTrp 364  
DB 241 ATCTCTCATGCTGTGGCAGAGCTCTCTGAGATTGGAGACCCGAGTTTACAGGAGCTG 300  
QY 364 TrpAsnSerGluSerValThrTyrPheTrpGlnAsnTrpAsnIleProValHisLysTrp 384  
DB 301 GTGGAATGCTGAGTCTGTCACTACTTTTGGCAGAACTGGAATATCCCGTGCAAGAATG 360  
QY 384 CyLeuLeuArgHisPheTyrLysProMetLeuArgGlySerSer 399  
DB 361 GTGCATCAGACACTTCTCAAGCCTATGCTCAGACTGGGACGACAC 406

RESULT 13  
CB775155 391 bp mRNA linear EST 16-MAY-2003  
LOCUS AMGNND:TRXP1-00005-C5-A trxp1 (10556) Rattus norvegicus cDNA clone  
DEFINITION trxp1-00005-C5 5', mRNA sequence.  
ACCESSION CB775155  
VERSION CB775155.1 GI:29863546  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Muridae; Rattus.  
REFERENCE 1 (bases 1 to 391)  
AUTHORS Angen EST Program.  
TITLE Angen Rat EST Program  
JOURNAL Unpublished (2003)  
COMMENT Contact: Dan Fitzpatrick  
Angen, Inc  
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00005 row: C column: 5.  
Location/Qualifiers  
1..391  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="trxp1-00005-C5"  
/tissue\_type="prostate"  
/clone\_lib="trxp1 (10556) "

## FEATURES

source

1..391  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="trxp1-00005-C5"  
/tissue\_type="prostate"  
/clone\_lib="trxp1 (10556) "

ORIGIN /note="Vector: pYVA-41L; Site\_1: HindIII; Site\_2: NotI;  
rat prostate"

Alignment Scores:  
Pred. No.: 8,736-62 Length: 391  
Score: 719.00 Matches: 127  
Percent Similarity: 100.0% Conservative: 3  
Best Local Similarity: 97.7% Mismatches: 0  
Query Match: 27.7% Indels: 0  
DB: 6 Gaps: 0

US-10-659-800-6 (1-488) x CB775155 (1-391)

QY 265 LeuAnpPheProArXSeProArXglLeArXglYsArXpHeuLeuArXglLeuGlu 284  
DB 1 CTCACCTTCTCCGATCCCGCAATACGAAAGCGCTTCTGACGGGGGCTTGTGAG 60  
QY 285 MetLeuPhePheThrGlnLeuGlnValGlyLeuIleGlnGlnTrpMetValProThrIle 304  
DB 61 ATGCTCTTTTACCCAGCTTCAAGTGGGGCTGATCCAGCATGATGATGCTCTACTT 120  
QY 305 GlnAsnSerMetLysProPheLysAspMetLysPheTyrSerArgIleIleGlnArgLeuLeu 324  
DB 121 CCAAGACTCCATGAGACCCCTTCAAGAGACATGAGATTCACGATCATTTGACGCTCTT 180  
QY 325 LysLeuAlaValProAsnHisLeuIleTrpLeuIlePhePheTyrTrpLeuPheHisSer 344  
DB 181 AAGCTGGGGGCTCCCAACCATGTGATAGGCTCATCTTCTTATGGCTTTTCCACTCA 240  
QY 345 CyLeuAlaValAlaGluLeuMetGlnPheGlnYsAspArgGluPheTyrArgAspTrp 364  
DB 241 ATCTCTCATGCTGTGGCAGAGCTCTCTGAGATTGGAGACCCGAGTTTACAGGAGCTG 300  
QY 365 TrpAsnSerGluSerValThrTyrPheTrpGlnAsnTrpAsnIleProValHisLysTrp 384  
DB 301 TGGGAATGCTGAGTCTGTCACTACTTTTGGCAGAACTGGAATATCCCGTGCAAGAATG 360  
QY 385 CyLeuLeuArgHisPheTyrLysProMetLeu 394  
DB 361 TGCAATCAGACACTTCTCAAGCCTATGCTC 390

RESULT 14  
AW391923 494 bp mRNA linear EST 04-FEB-2000  
LOCUS AW391923  
DEFINITION QV4-ST0233-251199-041-g07 ST0233 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW391923  
VERSION AW391923.1 GI:6896582  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 494)  
AUTHORS HCGP <http://www.judwig.org.br/ORGSTES>.  
TITLE The FAPESP/LICR Human Cancer Genome Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@judwig.org.br](mailto:asimpson@judwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.judwig.org.br/scripts/gethtml2.pl?l=QV4&t2=QV4-ST0233-251199-041-g07&t3=1999-11-25&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence stop: 494.  
Location/Qualifiers

## FEATURES

Location/Qualifiers

### Source

```

1. .494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="STR233"
/note="Organ: stomach; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

## ORIGIN

### Alignment Scores:

Pred. No.:	4,436-61	Length:	494
Score:	713.50	Matches:	134
Percent Similarity:	82.7%	Conservative:	0
Best Local Similarity:	82.7%	Mismatches:	2
Query Match:	27.5%	Indels:	27
DB:	1	Gaps:	1

US-10-659-800-6 (1-488) x AW391923 (1-494)

Oy	298	LIINTPmEctVaIProctHrlllegLIansSerMeLySProPheLySaPheLaSPYrSer	317
Db	487	CAGTGGATGGTCCCAACCATCCAGAACTCCATGAGACCTTCAAGACATGACTACTCA	428
Oy	318	ArgIIeIIegIuArgLeuLeuLySLeuAlaValProAsnHIsLeuIleTrpLeuIIephe	337
Db	427	CGCATCATCGAGTGCCTCCTGAAGCTGGCGGGTCCCAATCACCTCATCTGGCTCATCTTC	368
Oy	338	PheTrpTrpLeuPheHIsSerCySLeuAsnAlaValAlaGluLeuMetGlnPheGlyAap	357
Db	367	TTCTACTGGCTTTCACCTCTGCTGGCTGAATGCCGTGGCTGAGCTCATGCAATTGGAGAC	308
Oy	358	ArgGluPheTrpArgAspTrpTrpAsnSerGluSerValThrTrpPheTrpGlnAsnTrp	377
Db	307	CGGAGGTTCTACCGGGAGCTGGTGAACCTCGAGTCTGTCACTCATTTCTGGCAGAACTGG	248
Oy	378	AsnIleProValHIsLySLeuTrpCysIle-----	386
Db	247	AACATCCCTGTGCACAAAGTGGTCAAT-CAGGTAGCTGGGGTGTGTGTGTGTGATGTGG	189
Oy	387	-----ArgHIsPheTrpLyS	391
Db	188	AACATGGCTGTGAACCTGAACCGGCTTTTCATGCCCCCTCCTGCGAGACACTTCAAG	129
Oy	392	ProMetLeuAlaGArgLySLeuSerLySLeuTrpMetAlaArgThrGlyValPheLeuAlaSer	411
Db	128	CCCAATGCTTCGACGGGGCGAGCAGCAATGTGAATGGCAGGACAGGGGTGTTCCTGGGCTCG	69
Oy	412	AlaPhePheHIsGluTyuTrpLeuValSerValProLeuAlaGMePheArgLeuTrpAlaPhe	431
Db	68	GCTTCTTTCACAGATCACTGGTGAAGCGTCCCTGTGCAATGTTCGGCCTCTGGGGCTTC	9
Oy	432	ThrGly	433
Db	8	ACGGGC	3

## RESULT 15

BM857224

LOCUS	430 bp	mRNA	linear	EST 06-MAR-2002
BM57224				

DEFINITION	K-EST0141455	S21SNUS20	Homo sapiens	CDNA clone	S21SNUS20-78-D11 5'
LOCUS	BM037224	430 bp	INTRA	11881	201 00-1988-200

DEFINITION	MRNA sequence
K-ES10141455	

Accession BM857224 mRNA sequence.

ACCESSION BM857224  
 DEPOSITION 1  
 CT 10010000

VERSION BM857224.1 GI:19213

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

ORGANISM	Homo sapiens	Homo sapiens	Homo sapiens
ORGANISM	Homo sapiens	Homo sapiens	Homo sapiens

REFERENCE  
AUTHORS  
Hominidae; Homo.  
1 (bases 1 to 430)  
Kim, N.S., Hahn, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Chu, M. Y., Kim, M. R.

TITLE	21C Frontier Korean EST Project 2001
JOURNAL	Unpublished (2002)
COMMENT	Contact: Kim YS

## FEATURES

Source

1. 430  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S21SNM520-78-D11"  
/sex="F"  
/tissue\_type="Stomach"  
/cell\_type="Floating aggregates"  
/lab\_host="Top10P"  
/clone\_1ib="S21SNM520"  
/note="Organ: Stomach; Vector: pTZ19RP1; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then deapped  
with taabaco acid pyrophosphatase (TAP) and then deapped  
Intact mRNA was ligated with DNA-RNA linker including EcoRI  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dr-selected mRNA by  
priming with dr-tailed vector. The dr-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA stranded  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10P by electroporation method  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

## ORIGIN

**Alignment Scores:**

Pred. No.:	8.09e-61	Length:	430
Score:	710.00	Matches:	1418
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.3%	Mismatches:	0
Query Match:	27.4%	Indels:	0
DB:	3	Gaps:	0

US-10-659-800-6 (1-488) X BM857224 (1-430)

Oy	76	PhSeSerSerAapSerGlyPhSeSeramYrtrGrGlyLeuLeuAantPyGValValMet	95
Db	3	TTAGGTCTGACAGTGGCTTCAGCAACTACCGTGGCATCTCGAATCTGTGTGTGGATG	62
Oy	96	LeuLeleSeSerAapAahArgPheLeuGluuAanLeuLeuYrYrGlyLeleLeuVal	115
Db	63	CTATCTTAGCAATGCCGGTATTCTTCGAGAACTCATCAAGTATGGCATCTCGTG	122
Oy	116	AapProIleGlnValValSerLeuPheLeuYsaAapProHisSerTrpProAlaProCys	135
Db	123	GACCCCATCAAGTGGTCTTCCTGTCCTGGAAGATCCCTATACCTGGCCGCCCATGTC	182
Oy	136	LeuValIleAlaAlaAsnValPheAlaValAlaIlePheGlnValGluYrArgLeuAla	155
Db	183	CTGGTATATGGCGGCACATGCTCTTCTGTGGCTGCATTCACAGGTGAGAAAGCCCTGGCG	242
Oy	156	ValGlyAlaLeuThrGluGlnAlaGlyLeuLeuLeuLysValAlaAsnLeuAlaThrIle	175

```
Db 243 GTGGGTGCCCTGACGAGCAGCGGGGACTGCTGCACGTGGCCACCTGGCCACCAATT 302
QY 176 LeuGpPheProAlaIaValIleuLeuValGluSerIleThrProValGlySerLeu 195
Db 303 CTGTGTTTCCAGCGGCTGTGCTTACTGTTGAGTCTATCACTCCAGTGGGCTCCCTG 362
QY 196 LeuAlaLeuMetAlaHisThrIleLeuPheLeuIysLeuPheSerTyrArgAspValAsn 215
Db 363 CTGGCGCTGATGGCGCAGACATCCTCTTCTCAAGCTCTTCTCTACCGCGACGTCAAC 422
QY 216 SerTTP 217
Db 423 TCATGG 428
```

Search completed: May 7, 2006, 04:18:29  
Job time : 5792 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 6, 2006, 19:36:32 ; Search time 1363 Seconds  
(without alignments)  
3579.277 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MGDGSSRRRTGSRPSHG.....VLMYHYDVVINYEAPAAEA 488

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 7170814

Minimum DB seq length: 0  
Maximum DB seq length: 500

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-O=/abs/ABSSWB.epool/US10659800/runac\_05052006\_122305\_17180/app\_query.fasta\_1  
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.szlm500.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=500  
-HOST=abbe02p -USER=US10659800\_QCGN\_1\_1\_1929\_@runac\_05052006\_122305\_17180  
-NCPU=6 -ICPU=3 -NO\_MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_21.\*  
1: geneseqn1980s.\*  
2: geneseqn190s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: geneseqn2004as.\*  
13: geneseqn2004bs.\*  
14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	842	32.5	452	6	ABL181686 Human ova
2	644	24.8	450	6	AAD40443 Bovine di
3	612	23.6	371	6	ABL82916 Human ova
4	603.5	23.3	433	4	AAS41185 CDNA enco

5	603.5	23.3	433	4	ABA06648 Human CDN
6	603.5	23.3	433	6	ABV83985 Human pol
7	567	21.9	447	9	ACH42044 Human foe
8	332	12.8	471	9	ACH14767 Human adu
9	330.5	12.7	380	6	AB085322 Arabidops
10	232	8.9	299	6	Aaz45382 DNA encod
11	232	8.9	299	3	AaA88845 Mouse acy
12	222.5	8.6	400	13	ADU81175 Ryegrass
13	217	8.4	421	12	ACH92522 Human gen
14	199	7.7	400	6	AAD40442 Bovine di
15	197.5	7.6	381	3	Aaz45387 Acyl-CoA:
16	197.5	7.6	381	3	AaA88836 C. elegan
17	191.5	7.4	393	9	ACH46527 Human inf
18	191.5	7.4	399	9	ACH46069 Human inf
19	191.5	7.4	401	9	ACH45947 Human inf
20	191.5	7.4	418	9	ACH45946 Human inf
21	188.5	7.3	401	9	ACH46143 Human inf
22	187.5	7.2	401	9	ACH46631 Human inf
23	181	7.0	470	6	ABL93444 Arabidops
24	177.5	6.8	365	9	ACH47212 Human inf
25	173.5	6.7	402	9	ACH46232 Human inf
26	171.5	6.6	378	13	ADS54498 Bacterial
27	170	6.6	275	3	Aaz45374 Acyl-CoA:
28	170	6.6	275	3	AaA88840 Soybean a
29	170	6.6	275	5	AAS01313 Soybean s
30	164.5	6.3	383	9	ACH47844 Human inf
31	164	6.3	325	3	AAA10204 Rat liver
32	159	6.1	421	6	ABK29584 Colon ade
33	155	6.0	253	3	Aaz45376 Acyl-CoA:
34	155	6.0	253	3	AaA88849 Maize acy
35	155	6.0	253	5	AAS01315 Maize ste
36	148	5.7	357	12	ADO55089 Gene #186
37	143.5	5.5	386	13	ADU14134 Solid tum
38	140.5	5.4	258	13	ADU81173 Ryegrass
39	130.5	5.0	408	9	ACH46887 Human inf
40	127	4.9	267	3	Aaz45373 Acyl-CoA:
41	127	4.9	267	3	AaA88841 Soybean a
42	127	4.9	267	5	AAS01312 Soybean s
43	123.5	4.8	234	3	Aaz45372 Acyl-CoA:
44	123.5	4.8	234	3	AaA88839 Soybean a
45	123.5	4.8	234	5	AAS01311 Soybean s

## ALIGNMENTS

RESULT 1	ABL181686	standard; CDNA; 452 BP.
ID	ABL181686;	
AC	ABL181686;	
DT	17-MAY-2002 (first entry)	
DE	Human ovarian cancer related CDNA clone SEQ ID NO:4664.	
OS	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.	
XX	Homo sapiens.	
XX	WO200192581-A2.	
XX	06-DEC-2001.	
XX	29-MAY-2001; 2001WO-US017756.	
XX	26-MAY-2000; 2000US-0207484P.	
XX	(CORI-) CORIXA CORP.	
XX	Algate PA, Harlocker SL, Jones R;	
XX	WPI, 2002-122075/16.	

Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide.

Claim 1: SEQ ID NO 4664; 489pp; English.

The present invention describes a composition (I) comprising: carriers and immunostimulants; and a polypeptide (II) of a ovarian tumour polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1) from the 10912 nucleotide sequences as given in AB177023 to AB187934, (III) encoding (II) having a sequence (S2), a T cell population of (II), or antigen presenting cells that express (II). (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to (S1) can be used for detecting ovarian cancer in a patient's biological sample preferably serum or ovarian tissue. The method comprises contacting a biological sample from a patient with (IV), detecting the amount of polynucleotide hybridising to (IV) and comparing the amount to a predetermined cutoff value and thereby detecting ovarian cancer in the patient, where the amount of polynucleotide hybridising to (IV) is detected preferably by polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is useful for stimulating and/or expanding T cells specific for an ovarian tumour protein comprising contacting T cells with (III) or (II). (II) is useful in design and preparation of ribozyme molecules for inhibiting expression of the tumour polypeptides and proteins in tumour cells; and to isolate a full length gene from a suitable library e.g., a tumour cDNA library using well known techniques

Sequence 452 BP, 86 A, 151 C, 107 G, 108 T, 0 U, 0 Other;

#### Alignment Scores:

Pred. No.:	1,246-70	Length:	452
Score:	842.00	Matches:	150
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	32.5%	Indels:	0
DB:	6	Gaps:	0

US-10-659-800-6 (1-488) x AB181686 (1-452)

OY	253	TYR	TYR	PHE	LEU	PHE	ALA	PRO	THR	LEU	CYS	TYR	GLU	LEU	ASN	PHE	PRO	ARG	SER	PRO	ARG	272
DB	3	TACT	ACTT	CTCT	TGCG	CCCA	CTTG	GCTAC	GAG	CTCA	CACTT	TCC	CGCT	CTCC	CCGC	CC	62					
OY	273	ILE	ARG	VAL	ARG	PHE	LEU	VAL	ARG	ILE	GLU	VAL	ARG	LEU	VAL	ARG	ILE	GLU	VAL	ARG	ILE	292
DB	63	ATCC	GAA	GGCT	TTCT	CTG	CGAC	GGAT	CTT	GAG	ATG	CTG	TTCA	CCCA	GGCT	CCAG	122					
OY	293	VAL	GLY	LEU	ILE	GLN	INT	RP	ME	VAL	PRO	THR	ILE	GLN	ASN	SER	ME	LYS	PRO	PHEN	312	
DB	123	GTTG	GGCT	GAT	CCAG	AGAT	GGAT	GGAT	CCCA	CCAT	CCNA	AACT	CCCA	TGAA	AGCC	TTCA	AG	182				
OY	313	ASP	ME	ASP	TYR	SER	ARG	ILE	ILE	GLU	VAL	ARG	LEU	VAL	ARG	ILE	GLU	VAL	ARG	ILE	332	
DB	183	GAC	ATG	GACT	ACT	ACT	CGCAT	CA	TGCG	CCCT	CTCG	AACT	CGCG	GTCC	CCCA	TAC	CACT	CC	242			
OY	333	ILE	TRP	LEU	ILE	PHE	PHE	TYR	TRP	LEU	PHE	ILE	ASN	SER	CYS	LEU	ASN	ALA	VAL	ALA	352	
DB	243	ATCT	GCT	CTAT	CTT	CTT	CTA	CTG	GCT	CTT	CTC	CACT	CTCT	GCTA	ATG	CCGT	GCTG	CTG	ACT	302		
OY	353	ME	GLN	PHE	GLY	ASP	ARG	ILE	TRP	ARG	ASP	TRP	TRP	ASN	SER	GLY	SER	VAL	THR	TYR	372	
DB	303	ATG	CA	GTTT	GGAG	ACCG	GAG	ATT	CTAC	CCGG	GA	CTGG	TGA	AACT	CCG	AGT	CTG	CA	CTT	CA	362	
OY	373	PHE	TRP	GLN	ASP	TRP	ASN	ILE	PRO	VAL	ILE	SLY	TRP	CYS	ILE	ARG	ILE	ARG	ILE	ARG	392	
DB	363	TTCT	GGC	AGAA	CTG	GAAC	ATCC	CTG	GCAC	AA	GTG	TGAT	GCAT	GAC	CACTT	CTT	CA	CA	AGCC	422		
OY	393	ME	LEU	ARG	ARG	GLY	SER	SER	LYS	TRP	ME	402										
DB	423	ATG	CTT	CGAC	GGGG	CGAC	GAC	CAAG	TGG	ATG	452											

RESULT 2

AAD40443	ID	AAD40443	standard; cDNA; 450 BP.
XX	AC	AAD40443;	
XX	DT	22-OCT-2002	(first entry)
XX	DE	Bovine diacylglycerol acyltransferase (DGAT1) cDNA fragment.	
XX	KW	Bovine; diacylglycerol acyltransferase; genotyping; milk production; DGAT1; polymorphism; farming industry; transgenic; enzyme; ss.	
XX	OS	Bos taurus.	
XX	FT	Key	Location/Qualifiers
XX	FT	primer_bind	complement(85..104)
XX	FT	primer_bind	/tag= a
XX	FT	primer_bind	/bound_moiety= "DGAT1forRT66"
XX	FT	primer_bind	complement(144..156)
XX	FT	primer_bind	/tag= b
XX	FT	primer_bind	/bound_moiety= "DGAT1forRT1ess66"
XX	FT	primer_bind	complement(190..202)
XX	FT	primer_bind	/tag= c
XX	FT	primer_bind	/bound_moiety= "DGAT1with66(FAM)"
XX	FT	primer_bind	complement(223..235)
XX	FT	primer_bind	/tag= d
XX	FT	primer_bind	/bound_moiety= "Primer"
XX	FT	primer_bind	285..301
XX	FT	primer_bind	/tag= e
XX	FT	primer_bind	/bound_moiety= "DGAT1ess66(VIC)"
XX	PN	WO200236824-A1.	
XX	XX	10-MAY-2002.	
XX	PD	31-OCT-2001; 2001WO-NZ000245.	
XX	PF	31-OCT-2000; 2000NZ-00507888.	
XX	PR	06-DEC-2000; 2000NZ-00508652.	
XX	PA	(GEOR/ GEORGES M A J.	
XX	PI	(COPP/ COPPIERS W H R.	
XX	PI	(GRIS/ GRISART B M J.	
XX	PA	(SNEL/ SNELL R G.	
XX	PA	(REID/ REID S J.	
XX	PA	(FORD/ FORD C A.	
XX	PA	(SPEL/ SPELMAN R J.	
XX	PI	Georges MJ, Coppieters WH, Grisart BM, Snell RG, Reid SJ;	
XX	PI	Ford CA, Spelman RJ,	
XX	DR	WPI; 2002-500128/53.	
XX	XX	Determining genetic merit of a bovine with respect to milk composition	
XX	PT	and volume for improved milk production, comprises determining the	
XX	PT	diacylglycerol acyltransferase gene genotypic state of the bovine.	
XX	PS	Disclosure; Page 41-42; 128pp; English;	
XX	CC	The invention relates to a method of genotyping bovine for improved milk	
XX	CC	production traits which comprises determining the diacylglycerol	
XX	CC	acyltransferase (DGAT1) genotypic state of the bovine, wherein the DGAT1	
XX	CC	gene and polymorphisms have been found to be associated with such	
XX	CC	improved milk production traits. The method is useful for selecting a	
XX	CC	bovine having a desired DGAT1 genotypic state. It is also useful for the	
XX	CC	identification and selection of a bovine having one of the polymorphisms	
XX	CC	in its DGAT1 gene. Milk produced from selected bovine which is useful for	
XX	CC	making a dairy product provides a beneficial health effect. An antibody	
XX	CC	to the protein having DGAT1 activity is useful for inhibiting the	
XX	CC	activity of bovine DGAT1 in a lactating bovine so as to modulate milk	
XX	CC	production and/or milk solids content. DGAT1 nucleic acid and its	
XX	CC	fragments are useful in the farming industry. They are also useful to	
XX	CC	generate transgenic animals which are useful to investigate the molecular	

CC basis of DGAT1 action and to test a substance for the ability to prevent, CC slow or enhance DGAT1 activity. The present sequence is bovine DGAT1 cDNA CC fragment used to illustrate the method of the invention

XX Sequence 450 BP; 77 A; 165 C; 117 G; 91 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	8,876-52	Length:	450
Score:	644.00	Matches:	126
Percent Similarity:	89.3%	Conservative:	7
Best Local Similarity:	84.6%	Mismatches:	12
Query Match:	24.8%	Indels:	4
DB:	6	Gaps:	1

US-10-659-800-6 (1-488) x AAD40443 (1-450)

```
QY 182 ValValLeuLeuValGluSerIleThrProValGlySerLeuLeuAlaMetAlaHis 201
    |||
DB 3 GTGGCCTTCTCTCGAGCTATCACTCCAGTGGGCTCGTGTGCTGCTGCTGCTGCTAC 62
QY 202 ThTtLeuPheLeuLeuLeuPheSerTyrrArgAspValAsnSerTTrCyArgArg--- 220
    |||
DB 63 ACCATCTCTTCTCTCAAGCTGTCTCTCAAGGAGCTCAACCTGTGTGCTGAGAGAGGC 122
QY 221 -----AlaArgAlaValAlaAlaSerAlaGlyIleValSerSerAlaAla 237
    |||
DB 123 AGGCTGGGGCCCAAGGCTGCTTGGCAGGTAAGAGCCCAAGGGGAGCTGCC 182
QY 238 ProHsthrValSerTyrrProAspAsnLeuThrTyrrArgAspLeuTyrrPheLeuPhe 257
    |||
DB 183 CAGGCACTGAGAGTCAAGGCAAGCACTGAGCTCAAGCTCAAGCTCAAGCTCAAGCTCA 242
QY 258 AlaProThrIleuValGluLeuAsnProAspSerProArgIleArgValArgPhe 277
    |||
DB 243 GCCCCACCTGTGTAGAGCTCAACTTCCCGGCTCCCGGCTCCCGGCTCCGAAAGGCTTC 302
QY 278 LeuLeuArgArgIleLeuGluMetLeuPhePheThrGlnLeuGlnValGlyLeuIleGln 297
    |||
DB 303 CTGCTGGCGGCACTCTGAGATGCTGTCTTCCCAAGCTCCAGGTGGGGCTGATCCAG 362
QY 298 GlnTrpMetValProThrIleGlnAsnSerMetIlePheLeuAspMetAspTyrrSer 317
    |||
DB 363 CAGTGAATGATCCGCGGCAATCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 422
QY 318 ArgIleIleGluArgLeuLeuLeu 326
    |||
DB 423 CGCATCTGTGAGCGCTCTGAAAGCTG 449
```

## RESULT 3

```
ID ABL82916 standard; cDNA, 371 BP.
XX
AC ABL82916;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:5894.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
```

XX WPI, 2002-122075/16.

PT Composition for therapy and diagnosis of ovarian cancer comprising  
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
PT polypeptide, antibody specific to polypeptide or T cell expressing  
PT polypeptide.

PS Claim 1, SEQ ID NO 5894; 489pp; English.

CC The present invention describes a composition (I) comprising: carriers  
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)  
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL7934,  
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),  
CC or antigen presenting cells that express (II). (I) has cytostatic  
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for  
CC detecting ovarian cancer in a patient's biological sample preferably  
CC serum or ovarian tissue. The method comprises contacting a biological  
CC sample from a patient with (IV), detecting the amount of polynucleotide  
CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
CC value and thereby detecting ovarian cancer in the patient, where the  
CC amount of polynucleotide hybridising to (IV) is detected preferably by  
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
CC useful for stimulating and/or expanding T cells specific for an ovarian  
CC tumour protein comprising contacting T cells with (III) or (II). (III) is  
CC useful in design and preparation of ribozyme molecules for inhibiting  
CC expression of the tumour polypeptides and proteins in tumour cells; and  
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
XX library using well known techniques

SO Sequence 371 BP; 65 A; 136 C; 89 G; 81 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	7,48e-49	Length:	371
Score:	612.00	Matches:	123
Percent Similarity:	99.2%	Conservative:	0
Best Local Similarity:	99.2%	Mismatches:	1
Query Match:	23.6%	Indels:	1
DB:	6	Gaps:	0

US-10-659-800-6 (1-488) x ABL82916 (1-371)

```
QY 186 ValGluSerIleThrProValGlySerLeuLeuAlaMetAlaHsthrIleuPhe 205
    |||
DB 1 GTGAGTATCACTCCAGTGGCTCCCTGCGGCTGATGCGC-CACACCATCTCTTC 59
QY 206 LeuIleuPheSerTyrrArgAspValAsnSerTTrCyArgArgAlaArgAla 225
    |||
DB 60 CTCAAGCTCTCTCTCAAGGAGTCAACTCATGTGTCGCGCAGGCGCAAGGCT 119
QY 226 AlaSerAlaGlyIleValSerSerAlaAlaProHsthrValSerTyrrProAsp 245
    |||
DB 120 GCTCTGAGGAGGAAGGCGCAGTGTCTCTCCCGCACACCGTGAAGCTACCGGAC 179
QY 246 AsnLeuThrTyrrArgAspLeuTyrrPheLeuPheAlaProThrLeuCyrrGluLeu 265
    |||
DB 180 AATCTGACCTTACCGGAGCTTACTCTCTTCCGCCCCACCTGTGTGCTACGAGCTC 239
QY 266 AsnProArgSerProArgIleArgIleValArgPheLeuLeuArgArgIleLeuGluMet 285
    |||
DB 240 AACTTCCCGGCTCTCCCGGATCCGGAAGGCTTTCGTGCGAGCAGATCTTGAAGTG 299
QY 286 LeuPhePheThrGlnLeuGlnValGlyLeuIleGlnIleTrpMetValProThrIleGln 305
    |||
DB 300 CTGTTCTTCAAGCTCCAGTGGGCTGATCCAGAGTGAATGATGCCCATCATCCAG 359
QY 306 AsnSerMetIle 309
    |||
DB 360 AACTCAGGAAG 371
```

## RESULT 4

AAS41185

ID AAS41185 standard; cDNA; 433 BP.  
XX AAS41185;  
AC  
XX 17-DEC-2001 (first entry)  
DT  
XX  
DE cDNA encoding novel human enzyme polypeptide #401.  
XX  
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW lysase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
KW anti arthritic; nephrotropic; anticoagulant; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200155301-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001239.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216802P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0228343P.  
PR 01-SEP-2000; 2000US-0228344P.  
PR 01-SEP-2000; 2000US-0228345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 02-OCT-2000; 2000US-0237049P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246539P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.





PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0228343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233403P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234979P.  
PR 25-SEP-2000; 2000US-0234984P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 01-NOV-2000; 2000US-0244674P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251865P.  
PR 08-DEC-2000; 2000US-0251866P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-476161/51.  
XX DR P-PSDB; ABB10426.  
XX  
XX Isolated nucleic acid molecule encoding an inflammation-associated  
PT polypeptide is used in preventing, treating or ameliorating a medical  
PT condition.  
XX  
XX  
XX Claim 1; SEQ ID NO 314; 859pp + Sequence Listing; English.  
PS  
XX The present invention provides human cDNAs, proteins and related genomic  
CC DNA. These can be used in the treatment of neural, immune system,  
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
CC renal and proliferative disorders and inflammation. The present sequence  
CC is a cDNA of the invention  
XX  
XX  
SQ Sequence 433 BP; 87 A; 128 C; 111 G; 103 T; 0 U; 4 Other;  
Alignment Scores:  
Pred. No.: 6,04e-48 Length: 433  
Score: 603.50 Matches: 126  
Percent Similarity: 87.2% Conserves: 3  
Best Local Similarity: 85.1% Mismatches: 12  
Query Match: 23.3% Indels: 7  
DB: 4 Gaps: 1  
US-10-659-800-6 (1-488) x ABA06648 (1-433)  
QY 264 GluLeuAenPheProArgSerProArgIleArgIysaRgPheLeuLeuArgArgIleLeu 283  
DB 1 GAGCTCAACTTCCCGCTCCCGCATCCGGAAGCCCTTCTGCTCCGACGATCCTT 60  
QY 284 GluMetLeuPhePhePheThGlnLeuGlnValGlyLeuIleGlnIntPMeValProThr 303  
DB 61 GAGATGCTGTCTTCACCACTCCAGGTGGGCTGATCCAGCAGTGATGTCCTCCACC 120  
QY 304 IleGlnAnSerMetLysProPheLysAspMetAspTrpYrsArgYlleIleGluArgLeu 323  
DB 121 ATCCAGAACTCCATGAAGCCCTTCAAGACATGACTACTCAGCATCATCATGAGCGCTC 180

```

Qy 324 LeuLysLeuAlaValProAsnHisLeuIleTyrPhePheTyrTrpLeuPheHis 343
Db 181 CTGAAGCTGGGGGCTCCCAATCACTCATCTGGCTCATCTTCTTACTGGCTTCCAC 240
Qy 344 SerCysLeuAsnAlaValAlaGluLeuMetGlnPhe-GlyAspArgGluPheTyrArgAs 363
Db 241 TCTGTCTCGAATAGCCGTGGCTGACCTCATGACGATTGGGAGACCCGGAGTTCTACCGGGA 300
Qy 363 pTTPTrpAsnSerGluSerValThrTyrPhe---TrpGlnAsnTrpAsnIleProVal 381
Db 301 CTGGGGGAATTCACAGTGTGTTCACCTAATTCTGGGACAGAACCTGCTGTG 360
Qy 382 -HisLysTrpCysIleArg-HisPheTyrLysPProMetLeuArgGlySerSerLys 400
Db 361 GCACAGTGTGTGATTCATGACCAATTTTAAAGCCCAAGTTTCAGMGGGGACAGACGAA 420
Qy 401 TrpMetAlaArg 404
Db 421 TTGATGNCAGG 432

RESULT 6
ABV83985
ID ABV83985 standard; cDNA; 433 BP.
XX
AC ABV83985;
XX
DT 09-DEC-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 314.
XX
KW Human; noctropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischizy; antianaemic; antidiabetic; cancer;
KW antitumour; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW gene; BG.
XX
OS Homo sapiens.
XX
PN US2002090672-A1.
XX
PD 11-JUL-2002.
XX
PF 17-JAN-2001; 2001US-00764853.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0228343P.
PR 01-SEP-2000; 2000US-0228344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

```

```

PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.

XX
PA (ROSE/) ROSEN C A.
PA (RUBEN/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI: 2002-681727/73.
DR P-FSDB; ABP67013.
XX
PT Novel polypeptide useful for diagnosis, prognosis, prevention, and
PT treatment of immune, hyperproliferative, renal, respiratory,
PT cardiovascular, reproductive, endocrine, gastrointestinal and
PT neurological disorders.
XX
PS Claim 1; SEQ ID NO 314; 369pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABV83682-ABV84101) and proteins
CC (ABP66710-ABP67129) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 433 BP; 87 A; 128 C; 111 G; 103 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 6,04e-48 Length: 433
Score: 603.50 Matches: 126
Percent Similarity: 87.2% Conservative: 3
Best Local Similarity: 85.1% Mismatches: 12
Query Match: 23.3% Indels: 7
DB: 6 Gaps: 1

US-10-659-800-6 (1-488) x ABV83985 (1-433)
Qy 264 GluLeuAsnPheProArgSerProArgIleArgLysArgPheLeuArgArgIleLeu 283

```

Db 1 GAGCTCAACTTTTCCCGCTCTCCCGCATCCGAGACGGCTTTCTGCTGCGACGGATCTT 60

Qy 284 GIumEtLeuPhePheThrGlnLeuGlnValGlyLeuIleGlnIleTrpMetValProthr 303

Db 61 GAGATGCTGTTCTTTCACCCAGCTCCAGGTGGGGCTCATCGACAGTGGATGCCACACC 120

Qy 304 IlegInaAsnSerMetLysProPheLysAspMetAspTyrSerArgIleIleGluArgLeu 323

Db 121 ATCCAGAACTCCATGAAAGCCCTTCAAGGACATGATGACTACAGCATCATCGAGCCCTTC 180

Qy 324 LeuLysLeuAlaValProAsnHisLeuIleTrpLeuIlePhePheTyrTrpLeuPheHis 343

Db 181 CTGAAGCTGGCGCGCTCCCATCATCCATCTGGCTCATCTTCTTCACTGGGCTTCCAC 240

Qy 344 SerCybLeuAsnAlaValIleGluLeuMetGlnPheGlyAspArgGluPheTyrArgAs 363

Db 241 TCCTGCTGTAAGAAGCCGTGGCTGAGCTCAATGACATTTGGAGAGACGGGAGTTCTAACGGGA 300

Qy 363 PTrpTrpAsnSerGlnSerValThrTyrPhe---TrpGlnAsnTrpAsnIleProVal 381

Db 301 CTGGTGGGAAATTCGAGATGTGTTCTACCTAAATTTCTGGGCGAGACTGGAAACATCCCTGTG 360

Qy 382 HisLysTrpCybIleArgHisPheTyrLysProMetLeuArgArgLysSerSerLys 400

Db 361 GCACAAGTGTGCATTCAGACAATTTTAAAGCCCAKGTTCGAMGGGGGGCAGACGAA 420

Qy 401 TrpMetAlaArg 404

Db 421 TTGAATGAGCAGG 432

XX	RESULT 7
XX	ACH42044
XX	ID ACH42044 strand; cDNA; 447 BP.
XX	ACH42044;
XX	13-OCT-2003 (first entry)
XX	Human foetal brain cDNA #3411.
XX	Human; SS; sequencing by hybridisation; SMH; expressed sequence tag; EST;
KW	genome mapping; biodiversity; genetic disorder.
XX	
OS	Homo sapiens.
PN	US2003073623-A1.
XX	
PD	17-APR-2003.
XX	
PF	30-JUL-2001; 2001US-00918995.
XX	
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R T.
PA	(LABA/) LABAT I.
PA	(STAC/) STACHE-CRAIN B.
PA	(DICK/) DICKSON M C.
PA	(JONE/) JONES L W.
XX	
PI	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX	
DR	WPI; 2003-615964/58.
XX	
PT	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
XX	
PS	Claim 1; SEQ ID NO 29256; 44bp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was

CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPO at  
CC [seqdata.uspto.gov/sequence.html?DocID=20030073623](http://seqdata.uspto.gov/sequence.html?DocID=20030073623)

Sequence 447 BP; 91 A; 124 C; 130 G; 100 T; 0 U; 2 Other;

**Alignment Scores:**

Pred. No.:	1.9e-44	Length:	447
Score:	567.00	Matches:	107

Percent Similarity:	77.7%	Conservative:	1
Best Local Similarity:	77.0%	Mismatches:	0

Query Match:	21.9%	Indels:	31
DB:	9	Gaps:	2

US-10-659-800-6 (1-488) X ACH42044 (1-447)

Qy 298 GlnTrp-----MetValProThrIleGlnAsnSerMetLysProphelys----- 312

Db 31 CGGTGAATTCGCGGATGTCCCCACCATCAGAACTCCATGAAGCCCTTCAAGGGAGT 90

Qy 312 ----- 312

Db 91 GGCTCATGTGCTCTTGACAGCTGGGGTGGCTGGGGAGTGACCAAGAGCATGGCTAGCTGAA 150

QY 313 -----AspMetAspTyrSerArgIleIleGluArgLeuLeuLysLeuAl 327

Db 151 GGGCTTGTTCGACGACATGGA~~CTACT~~ACGCATCATCGAGCGCCTCCTGAAGCTGCC 210

327 aValProAsnIaLeuIeTrpLeuIePhePheTyTrpLeuPheIaSerCysLeuAs 347

Db 211 GGTCCCAATCAGCTCATCTGGCTCATCTTCTTCTACTGGCTCTTCCACTCCCTGAA 270

QY 347 **NA**IAVA**LA**AGLU**LE**Me**TC**IN**Ph**e**GL**YA**PP**arg**GL**U**Ph**e**TY**ARG**AS**PT**PT**TP**AS**NS**E** 367

Db 271 TGCCGTGGCTGAGCTCATGCA GTTGGAGACCGGAGTCTTACCGGACTGGTGA ACTC 330

QY 367 rgluSerValThrTyrPheTrpGlnAsnTrpAsnIleProValHisLysTrpCysIleAr 387

Db 331 CGAGTCTGTCACTACTTCTGGCAGACTGGACATCCCTGTGCACAAGTGTGCATCAG 390

QY 387 gHISpHeTYrLYsPRoMeTLLeuARGgLYSeRSeLYsTRMeTAlAArgThr 405

Db 391 ACACTTCTACAAGCCCATGCTTCGACGGGGCAGCAGCAAGTGATGGCCAGGACA 445

RESULT 8  
10/11/2007

ID ACH14767 standard; cDNA; 471 BP.

AC ACH14767;  
vv

DT 13-OCT-2003 (first entry)  
vv

DE Human adult brain cDNA #1979.

KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST; genome mapping; biodiversity; genetic disorder

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

XX  
PN  
11C3003073633-A1





PF 12-APR-2000; 2000MO-US009696.  
XX  
XX 12-APR-1999; 99US-0128995P.  
XX  
XX (MONS ) MONSANTO CO.  
XX  
XX Venkatramesh M, Corbin DR, Bhat GB, Boddupalli SS, Grebenok RJ,  
PI Kishore GM, Lardizabal KD, Lassner MW, Rangwala SH, Karunanandaa B;  
XX WPI; 2000-65136/64.  
XX  
XX  
XX Genetically engineering the biosynthetic pathways in plants involved in  
PT the accumulation of sterol compounds and tocopherol to produce compounds  
PT for lowering the level of low density lipoprotein cholesterol in blood  
PT serum.  
XX  
XX  
XX Disclosure; Page 62-63; 166pp; English.  
XX  
XX The present sequence is that of mouse acyl CoA:cholesterol  
CC acyltransferase (ACAT) expressed sequence tag (EST) DNA sequence 11.  
CC Mouse ACAT EST DNA sequence 1 is given in AAA88844. Sterol O-  
CC acyltransferases such as ACAT catalyse the formation of cholesterol  
CC esters from cholesterol and long chain fatty acids. Recombinant  
CC constructs of the invention are used to alter the biosynthesis and  
CC accumulation of sterols and tocopherols in transgenic plants. Seeds of  
CC such plants may contain elevated levels of sitosterol and/or its esters,  
CC and alpha-tocopherol, and reduced levels of campesterol and campestanol  
CC and their esters. The seeds may also contain the novel sterol  
CC brassicasteranol. Oil obtained from the seeds can be used in food and  
CC pharmaceutical compositions to lower levels of low density lipoprotein  
CC cholesterol in blood serum. ACAT enzymes can be used in the present  
CC invention to produce elevated levels of phytoosterol and/or phytosterol  
CC esters  
XX  
XX  
SQ Sequence 299 BP; 50 A; 89 C; 85 G; 75 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 8.53e-13 Length: 299  
Score: 232.00 Matches: 44  
Percent Similarity: 96.1% Conservative: 5  
Best Local Similarity: 86.3% Mismatches: 2  
Query Match: 8.9% Indels: 1  
DB: 3 Gaps: 0  
  
US-10-659-800-6 (1-488) x AAA88845 (1-299)  
QY 434 MetMetAlaGlnIleProLeuAlaTrpPheValGlyArgPheGlnGlyAsnTyrGly 453  
Db 3 ATGATGGCTCAAGTCCCACTGGCTGGATTTGGGCGCGATTTCTCCAGGGAACCTATGGC 62  
QY 454 AsnAlaValAlaTrpLeuSerLeuIleIleGlyGlnProIleAlaValLeuMetTyrVal 473  
Db 63 AATGAGCTGTGTGGGTGACACTCATTCATCTGGGCAACCGGTGGCTGT-CTCATGTATGTC 121  
QY 474 HisAspTyrTyrValLeuAsnTyrGlnAlaPro 484  
Db 122 CACGACTACTACTGTCTCACTACGATGCGCCA 154  
  
RESULT 12  
ADU81175  
ID ADU81175 standard; cDNA; 400 BP.  
XX  
XX AC ADU81175;  
XX  
XX 10-FEB-2005 (first entry)  
XX  
XX Ryegrass diacylglycerol acyltransferase, DGAT, predicted cDNA.  
XX  
XX diacylglycerol acyltransferase; DGAT; plant breeding; crop improvement;  
XX fatty acid biosynthesis; ss.  
XX  
XX Lolium perenne.  
XX

PN WO2004101793-A1.  
XX  
XX PD 25-NOV-2004.  
XX  
XX PF 14-MAY-2004; 2004WO-AU000635.  
XX  
XX 16-MAY-2003; 2003AU-00902413.  
XX  
XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
XX  
XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
PI Bryan GT, Burling MF, Roberts NJ, Trollope AJ, Woodfield DR;  
XX WPI; 2004-821891/81.  
XX  
XX  
XX New nucleic acid from a ryegrass (Lolium) or fescue (Festuca) species  
PT encoding a diacylglycerol acyltransferase 1 (DGAT1) or DGAT1-like  
PT polypeptide, useful as a genetic marker or for modifying fatty acid  
PT biosynthesis in a plant.  
XX  
XX  
XX Example 1; SEQ ID NO 38; 188pp; English.  
XX  
XX  
XX The invention relates to a new substantially purified or isolated nucleic  
CC acid or its fragment from a ryegrass (Lolium) or fescue (Festuca) species  
CC encodes a diacylglycerol acyltransferase 1 (DGAT1) or DGAT1-like  
CC polypeptide or its fragment or variant. The nucleic acid or nucleic acid  
CC fragment and/or its nucleotide sequence information or single nucleotide  
CC polymorphism of perennial ryegrass Lolium perenne is useful as a  
CC molecular genetic marker or for modifying fatty acid biosynthesis in a  
CC plant. The present sequence represents a plant diacylglycerol  
CC acyltransferase, DGAT, cDNA.  
XX  
XX  
SQ Sequence 400 BP; 91 A; 85 C; 93 G; 131 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 1.05e-11 Length: 400  
Score: 222.50 Matches: 51  
Percent Similarity: 50.8% Conservative: 15  
Best Local Similarity: 39.2% Mismatches: 40  
Query Match: 8.6% Indels: 25  
DB: 13 Gaps: 3  
  
US-10-659-800-6 (1-488) x ADU81175 (1-400)  
QY 373 PheTrpGlnAsnTrpAsnIleProValHisIleTyrCyHisArgHisPheTyrIlePro 392  
Db 1 TACTGAGAGATGTGGAATATGCTGTGCTATATATGCTGTTCGCCATATATATTTCCC 60  
QY 393 MetLeuArgArgGlySerSerIleTrpMetAlaArgThrGlyValPheLeuAlaSerAla 412  
Db 61 CCCAGGCGAGTGTATATCAAGAGAGTGTCTGTTGTATCATTTTGTATCTGCGC 120  
QY 413 PhePheHisGlyTyrLeuValSerValProLeuArgMetPheArgLeuTrpAlaPheThr 432  
Db 121 GTGCTCCATGAGTTATGTGTGCTGTCTCCCGCAATGTGCAGAGTTCTGGGCATTTCTTA 180  
QY 433 GlyMetMetAlaGlnIleProLeu-----AlaTrpPheValGlyArgPhePhe 448  
Db 181 GGGATCATGTGCGAATCCCTTTTATCATATTGACATCATCTTAAGCAAAATTCAGA 240  
QY 449 GlnGlyAsnTyrGlyAsnAlaAlaValTrp----- 458  
Db 241 GATACATATGCGCGGCAATCATATATCTGTCTTTTCTGATCTAGGACAGGCTATG 300  
QY 459 -----LeuSerLeuIleIleGlyGlnProIleAlaValLeu 470  
Db 301 CCGGCAATGATATTCTGTCTTTCTTCTG-ATCTACGACAGCCTATGTGCTTTC 359  
QY 471 MetTyrValHisAspTyrTyrValLeuAsn 480  
Db 360 CTGTACTACCATGAT-----GTGATGAT 383  
  
RESULT 13

ACH92522/c  
 ID ACH92522 standard; DNA; 121 BP.  
 XX  
 AC ACH92522;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 DE Human genome derived single exon probe #25717.  
 XX  
 KW Human; probe; ss; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003194704-A1.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 03-APR-2002; 2002US-00029386.  
 XX  
 PR 03-APR-2002; 2002US-00029386.  
 XX  
 PA (PENN/) PENN S G.  
 PA (RANK/) RANK D R.  
 PA (HANZ/) HANZEL D K.  
 PI Penn SG, Rank DR, Hanzel DK;  
 XX  
 DR WPI; 2004-119264/12.  
 XX  
 PT New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 XX  
 XX  
 PS Claim 1; SEQ ID NO 25717; 80bp; English.  
 XX  
 CC The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridizes under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC a method of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subscription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above. The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterizing  
 CC alternative splicing events, in detecting and characterizing gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
 CC  
 XX

SQL Sequence 121 BP; 27 A; 30 C; 39 G; 25 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 6,14e-12 Length: 121  
 Score: 217.00 Matches: 40  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 8.4% Indels: 0  
 DB: 12 Gaps: 0  
 US-10-659-800-6 (1-488) x ACH92522 (1-121)  
 QY 445 GlyArgPhePheGinglyAsnTyrglyAsnAlaAlaValTyrPleuSerLeuIleIleGly 464  
 DB 121 GCGCGCTTTTCCAGGGCAACTATGCAACGACAGCTGTGGCTGTCCGTCATCATCCGA 62  
 QY 465 GlnProIleAlaValLeuMetTyrValHisAspTyrTyrValLeuAsnTyrGlnAlaPro 484  
 DB 61 CAGCCAAATGAGCGCTTCATGTATGCTCCAGCACTACGTCAGTCACTATGAGGCCCA 2  
 RESULT 14  
 AAD40442  
 ID AAD40442 standard; DNA; 400 BP.  
 AC AAD40442;  
 XX  
 DT 22-OCT-2002 (first entry)  
 XX  
 DE Bovine diacylglycerol acyltransferase (DGAT1) gene, Q allele.  
 XX  
 KW Bovine; diacylglycerol acyltransferase; genotyping; milk production;  
 KW DGAT1; polymorphism; farming industry; transgenic; enzyme; gene; de.  
 XX  
 OS Bos taurus.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT primer\_bind /tag= a  
 FT /bound\_moiety= "DGAT1.23"  
 FT complement(225..244)  
 FT primer\_bind /tag= b  
 FT /bound\_moiety= "DGAT1 11"  
 FT primer\_bind /tag= c  
 FT /bound\_moiety= "DGAT1 14B"  
 FT primer\_bind /tag= d  
 FT /bound\_moiety= "DGAT1 12"  
 XX  
 PN WO200236824-A1.  
 XX  
 PD 10-MAY-2002.  
 XX  
 PF 31-OCT-2001; 2001WO-NZ000245.  
 XX  
 PR 31-OCT-2000; 2000NZ-00507888.  
 XX  
 PR 06-DEC-2000; 2000NZ-00508652.  
 XX  
 PA (GBOR/) GEORGES M A J.  
 PA (COPP/) COPPIETERS W H R.  
 PA (GRIS/) GRISART B M J.  
 PA (SNEL/) SNEL R G.  
 PA (REID/) REID S J.  
 PA (FORD/) FORD C A.  
 PA (SPEL/) SPELMAN R J.  
 XX  
 PI Georges MAJ, Coppieters WHR, Grisart BMJ, Snell RG, Reid SJ;  
 PI Ford CA, Spelman RJ;  
 XX  
 DR WPI; 2002-500128/53.  
 XX  
 PT Determining genetic merit of a bovine with respect to milk composition  
 PT and volume for improved milk production, comprises determining the





```

Qy      101 AlaArgLeuPheLeuGluAsnLeuIleuSerTyrgIyileLeuValAspProileGlnVal 120
Db      251 GGAGCGGTGGCACTTGAAATGTGATCAATATGTAATTTGATATACACCCCTTCAGTGG 310
Qy      121 ValSerLeuPheLeuLysAspProH18-----SerTTPProAlaProCysLeuVal 137
Db      311 ATCTCAACGTTGTGTGACATCATCAATTTGAGCTGGCAAAATCTTGCTCTCATC 370
Qy      138 IleAlaIle 140
Db      371 CTATGCTCA 379

```

Search completed: May 6, 2006, 22:12:35  
 Job time : 1371 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search using frame.p2n model

Run on: May 7, 2006, 02:34:18 ; Search time 7600 Seconds  
(without alignments)  
5474.922 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MDRGSSRRRTGSRPSHG.....VLMYVDYVLYNEAPAAEA 488

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 6077154

Minimum DB seq length: 0  
Maximum DB seq length: 500

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame.p2n.model -DEV=xlh  
-O=/abs/ABSWEB.spool/US1065800/runat.05052006.122306.17192/app.query.fasta\_1  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.szlm500.rge -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=100 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=500  
-HOST=abes03h -USER=US1065800 @CGN 1.1 4939 @runat.05052006.122306.17192  
-NCPUB6 -ICPU=3 -NO MMAP -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:.\*  
1: gb\_da:.\*  
2: gb\_in:.\*  
3: gb\_env:.\*  
4: gb\_cm:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pr:.\*  
9: gb\_ro:.\*  
10: gb\_sts:.\*  
11: gb\_sv:.\*  
12: gb\_un:.\*  
13: gb\_vl:.\*  
14: gb\_hlg:.\*  
15: gb\_pl:.\*  
  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES  
Result No. Score Match Length DB ID Description  
-----  
1 842 32.5 452 6 CQ461886 CQ461886 Sequence  
2 612 23.6 371 6 CQ463116 CQ463116 Sequence  
3 367 14.1 201 10 BV207924 sqmm22493

4 352 13.6 201 10 BV207926 BV207926 sqmm22493  
5 232 8.9 299 6 BD241855 BD241855 Acyl CoA:  
6 232 8.9 299 6 AR227750 AR227750 Sequence  
7 232 8.9 299 6 AX090338 AX090338 Sequence  
8 197.5 7.6 381 6 BD241860 BD241860 Acyl CoA:  
9 197.5 7.6 381 6 AR227755 AR227755 Sequence  
10 176 6.8 113 6 BD118287 BD118287 EST and e  
11 176 6.8 113 6 AR422734 AR422734 Sequence  
12 176 6.8 113 6 AX983428 AX983428 Sequence  
13 170 6.6 275 6 BD241847 BD241847 Acyl CoA:  
14 170 6.6 275 6 AR227742 AR227742 Sequence  
15 170 6.6 275 6 AX090352 AX090352 Sequence  
16 168 6.5 310 10 BV102527 BV102527 MARC 1757  
17 164 6.3 225 6 AR121924 AR121924 Sequence  
18 164 6.3 225 6 BD223360 BD223360 Toxicolog  
19 159 6.1 421 6 AX379068 AX379068 Sequence  
20 155 6.0 253 6 BD241849 BD241849 Acyl CoA:  
21 155 6.0 253 6 AR227744 AR227744 Sequence  
22 155 6.0 253 6 AX090354 AX090354 Sequence  
23 143.5 5.5 386 6 CQ923373 CQ923373 Sequence  
24 127 4.9 267 6 BD241846 BD241846 Acyl CoA:  
25 127 4.9 267 6 AR227741 AR227741 Sequence  
26 127 4.9 267 6 AX090351 AX090351 Sequence  
27 123.5 4.8 234 6 BD241845 BD241845 Acyl CoA:  
28 123.5 4.8 234 6 AR227740 AR227740 Sequence  
29 123.5 4.8 234 6 AX090350 AX090350 Sequence  
30 103 4.0 200 6 AR672183 AR672183 Sequence  
31 101 3.9 65 6 CQ532056 CQ532056 Sequence  
32 100 3.9 65 6 CQ556741 CQ556741 Sequence  
33 97 3.7 129 8 F331502S12 F331513 Homo sapi  
34 97 3.7 440 10 BV422477 BV422477 S229P69FF  
35 94 3.6 430 6 CQ675847 CQ675847 Sequence  
36 91 3.5 471 6 CQ051475 CQ051475 Sequence  
37 91 3.5 471 6 CQ066524 CQ066524 Sequence  
38 91 3.5 471 6 CQ093572 CQ093572 Sequence  
39 91 3.5 471 6 CQ132342 CQ132342 Sequence  
40 91 3.5 471 6 CQ170912 CQ170912 Sequence  
41 91 3.5 471 6 CQ200059 CQ200059 Sequence  
42 91 3.5 471 6 CQ215540 CQ215540 Sequence  
43 91 3.5 471 6 CQ254134 CQ254134 Sequence  
44 91 3.5 471 6 CQ291147 CQ291147 Sequence  
45 91 3.5 471 6 CQ328262 CQ328262 Sequence

## ALIGNMENTS

RESULT 1  
LOCUS CQ461886 452 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 4664 from Patent WO0192581.  
ACCESSION CQ461886  
VERSION CQ461886.1 GI:41427505  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cnidaria; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1  
AUTHORS Algate,P.A., Harlocker,S.L. and Jones,R.  
TITLE Compositions and methods for the therapy and diagnosis of  
ovariancancer  
JOURNAL Patient. MO 0192581-A 4664 06-DEC-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..452  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores: 2.1e-54 Length: 452  
Pred. No.: 452



Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 14.1% Indels: 0  
DB: 10 Gaps: 0  
US-10-659-800-6 (1-488) x BV207924 (1-201)

QY 380 ProValHisLysLeuPheCysIleArgHisPheTyrLysProMetLeuAlaArgLysSer 399  
DB 3 CCGTGGCAACAAGTGGATGATGACACCTCTTCAAGCCATGCTTGCAGCGGCGACGAC 62

QY 400 LysTTPMetAlaArgThrGlyValPheLeuAlaSerAlaPhePheHisGluTyrLeuVal 419  
DB 63 AAGTGATGGCCAGGACAGGGGGTGTCTCGGCTCGGCTTCTTCCACAGATCCGATG 122

QY 420 SetValProLeuArgMetPheArgLeuTPAlaPheThrGlyMetMetAlaGlnIlePro 439  
DB 123 AGCGTCCCTCTCGAATGTTCCGCTCTGGCGTTTCACGGCATGATGGCTCAGATCCA 182

QY 440 LeuAlaTTPPheValGly 445  
DB 183 CTGGCTGGTTCGTGGC 200

RESULT 4  
BV207926 201 bp DNA linear STS 10-JUN-2004  
LOCUS sqm224934 Human DNA (Sequenom) Homo sapiens STS genomic, sequence  
DEFINITION tagged site.  
ACCESSION BV207926  
VERSION BV207926.1 GI:48178556  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 201)  
AUTHORS Nelson,R.M., Marellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
TITLE Cancer,C.R. and Braun,A.  
large-scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
JOURNAL Genome Res. (2004) In press  
COMMENT Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 201.

FEATURES  
source Location/Qualifiers  
1..201  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1..>201

ORIGIN  
STS  
Alignment Scores:  
Pred. No.: 4.56e-18 Length: 201  
Score: 352.00 Matches: 66  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 13.6% Indels: 0  
DB: 10 Gaps: 0  
US-10-659-800-6 (1-488) x BV207926 (1-201)

QY 398 SerSerLysTTPMetAlaArgThrGlyValPheLeuAlaSerAlaPhePheHisGluTyr 417  
|||||

DB 3 AGCAGCAAGTGATGGCCAGGACAGGGGTGTTCTGGCCTCGGCTTCTTCCACGATGAC 62

QY 418 LeuValSerValProLeuArgMetPheArgLeuTPAlaPheThrGlyMetMetAlaGln 437  
DB 63 CTGGTAGGCTCCCTCGAATGTTCCGCTCTGGCGCTTTCACGGCATGATGGCTCAG 122

QY 438 IleProLeuAlaTTPPheValGlyArgPhePheGlnGlyAsnTyrGlyAsnAlaVal 457  
DB 123 ATCCCACTGGCTGATGTTCTGGCGCGCTTTTCCAGGCACTATGCAACCAAGCTG 182

QY 458 TTPLeuSerLeuIleIle 463  
DB 183 TGGCTGCTGCTCATCATC 200

RESULT 5  
BD241855 299 bp DNA linear PAT 17-JUL-2003  
LOCUS BD241855  
DEFINITION Acyl CoA:cholesterol acyltransferase related nucleic acid  
sequences.  
ACCESSION BD241855  
VERSION BD241855.1 GI:33051625  
KEYWORDS JP 2002517201-A/12.  
SOURCE Mus sp.  
MUS SP.  
ORGANISM Mus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 299)  
AUTHORS Laesner,M.W. and Ruezinsky,D.M.  
TITLE Acyl CoA:cholesterol acyltransferase related nucleic acid sequences  
JOURNAL Patent: JP 2002517201-A 12 18-JUN-2002;  
CALGENE LLC

COMMENT OS Mus sp. (murine)  
PN JP 2002517201-A/12  
PD 18-JUN-2002  
PF 04-JUN-1999 JP 2000552290  
PR 05-JUN-1998 US 60/088143, 12-NOV-1998 US 60/108389 PI  
MICHAEL W LAESNER,DIANE M RUEZINSKY  
PC C12N15/09,A01H5/00,C11B1/00,C11C3/00,C12N1/21,C12N5/10,C12N9/  
PC 10,C12N15/00,  
PC C12N5/00  
CC Acyl CoA:cholesterol acyltransferase related nucleic acid CC

FEATURES  
FH Key Location/Qualifiers  
FT source 1..299  
FT Location/Qualifiers  
1..299  
/organism="Mus sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10095"

ORIGIN  
Alignment Scores:  
Pred. No.: 7.75e-09 Length: 299  
Score: 232.00 Matches: 44  
Percent Similarity: 96.1% Conservative: 5  
Best Local Similarity: 86.3% Mismatches: 2  
Query Match: 8.9% Indels: 1  
DB: 6 Gaps: 0  
US-10-659-800-6 (1-488) x BD241855 (1-299)

QY 434 MetMetAlaGlnIleProLeuAlaTTPPheValGlyArgPhePheGlnGlyAsnTyrGly 453  
DB 3 ATGATGGCTGAGTCCCGCTGGCTGATTTGTGGGCCGATTTCTTCAAGGAACTATGGC 62

QY 454 AsnAlaValATTPLeuSerLeuIleIleGlyGlnProIleAlaValLeuMetTyrVal 473  
DB 63 AATGCACTGTGTGGGTACACTCATCTTGGGCAACGGGTGGCTGT-CTCATGTATGTTC 121  
|||||

QY 474 HisAspTyrTyrValLeuAsnTyrGluAlaPro 484  
|||||

Db 122 CACGACTACTAGCTGCTCACTACGATGCCCA 154

RESULT 6 AR227750 AR227750 299 bp DNA linear PAT 20-DEC-2002

LOCUS AR227750

DEFINITION Sequence 13 from patent US 6444876.

ACCESSION AR227750

VERSION AR227750.1 GI:27266341

KEYWORDS

SOURCE

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 299)

AUTHORS Laessner,M.W. and Ruezinsky,D.M.

TITLE Acyl CoA: cholesterol acyltransferase related nucleic acid sequences

JOURNAL Patent: US 6444876-A 13 03-SEP-2002;

Calgene LLC; Davis, CA;

MOX:

FEATURES

source 1..299

Location/Qualifiers

ORIGIN

Alignment Scores:

Pred. No.: 7.75e-09 Length: 299

Score: 232.00 Matches: 44

Percent Similarity: 96.1% Conservative: 5

Best Local Similarity: 86.3% Mismatches: 2

Query Match: 8.9% Indels: 1

Gaps: 0

US-10-659-800-6 (1-488) x AR227750 (1-299)

Qy 434 MetMetAlaGlnIleProLeuAlaTrpPheValGlyArgPhePheGlnGlyAsnTyrGly 453

Db 3 ATGATGGCTCAGGCTCCCACTGGCTGATGTGGGCCGATCTTCCAGGGAACATGAGC 62

Qy 454 AsnAlaAlaValaTrpLeuSerLeuIleIleGlyGlnProIleAlaValLeuMetTyrVal 473

Db 63 AATGACGCTGTGGTGATCACTCATCTGAGCAACCGGTGGCTGT-CTCATGTATGTC 121

Qy 474 HisAepTyrTyrValLeuAsnTyrGluAlaPro 484

Db 122 CACGACTACTAGCTGCTCACTACGATGCCCA 154

RESULT 7

AX090338 AX090338 299 bp DNA linear PAT 21-MAR-2001

LOCUS AX090338

DEFINITION Sequence 31 from Patent WO0116308.

ACCESSION AX090338

VERSION AX090338.1 GI:13444204

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Laessner,M. and van Bennekom,A.

TITLE Plant sterol acyltransferases

JOURNAL Patent: WO 0116308-A 31 08-MAR-2001;

MONSANTO COMPANY (US)

FEATURES

source 1..299

Location/Qualifiers

ORIGIN

Alignment Scores:

Pred. No.: 7.75e-09 Length: 299

Score: 232.00 Matches: 44

Percent Similarity: 96.1% Conservative: 5

Best Local Similarity: 86.3% Mismatches: 2

Query Match: 8.9% Indels: 1

Gaps: 0

US-10-659-800-6 (1-488) x AX090338 (1-299)

Qy 434 MetMetAlaGlnIleProLeuAlaTrpPheValGlyArgPhePheGlnGlyAsnTyrGly 453

Db 3 ATGATGGCTCAGGCTCCCACTGGCTGATGTGGGCCGATCTTCCAGGGAACATGAGC 62

Qy 454 AsnAlaAlaValaTrpLeuSerLeuIleIleGlyGlnProIleAlaValLeuMetTyrVal 473

Db 63 AATGACGCTGTGGTGATCACTCATCTGAGCAACCGGTGGCTGT-CTCATGTATGTC 121

Qy 474 HisAepTyrTyrValLeuAsnTyrGluAlaPro 484

Db 122 CACGACTACTAGCTGCTCACTACGATGCCCA 154

RESULT 8

BD241860 BD241860 381 bp DNA linear PAT 17-JUL-2003

LOCUS BD241860

DEFINITION Acyl CoA:cholesterol acyltransferase related nucleic acid sequences.

ACCESSION BD241860

VERSION BD241860.1 GI:33051630

KEYWORDS JP 2002517201-A/17.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 381)

AUTHORS Laessner,M.W. and Ruezinsky,D.M.

TITLE Acyl CoA:cholesterol acyltransferase related nucleic acid sequences

JOURNAL Patent: JP 2002517201-A 17 18-JUN-2002;

CALGENE LLC

COMMENT OS Caenorhabditis elegans (nematode)

PN JP 2002517201-A/17

PD 18-JUN-2002

PF 04-JUN-1999 JP 2000552290

PR 05-JUN-1998 US 60/088143 12-NOV-1998 US 60/108389 PI

MICHAEL W LASSNER,DIANE M RUEZINSKY

PC C12N15/09,A01H5/00,C11B1/00,C11C3/00,C12N1/21,C12N5/10,C12N9/00

PC C12N5/00

CC n at position 46 is unknown

CC key Location/Qualifiers

FT misc\_feature (46).

Location/Qualifiers

source 1..381

Location/Qualifiers

ORIGIN

Alignment Scores:

Pred. No.: 4.11e-06 Length: 381

Score: 197.50 Matches: 41

Percent Similarity: 61.2% Conservative: 22

Best Local Similarity: 39.8% Mismatches: 31

Query Match: 7.6% Indels: 9

Gaps: 2

US-10-659-800-6 (1-488) x BD241860 (1-381)

Qy 47 ProAlaProAlaProAsn-----LysAapGlyAapAlaGlyValGly 60

Db 71 CCAAGACGCTCCCTCATTTGCACAAATGTAATTCCTCAAGGGAAGTTCAGAAATG 130

Qy 61 SerGlyHisTrpGluLeuArgCysHisArgLeuGlnAapSerLeuPheSerSerAapSer 80

Db 131 AGAGACCTTGCGAGAAAGTGATCATATCTGCTCAAGATTCATTTGTTTCAGACGATTC 190

Qy	81	GLYPHESerAntyArxGlyyleuEunentPcySvalValMeleuileleuSerAn	100
Db	191	GGATGCAAAATTTCCGTGGATCTTCCTCAATTTGCTATTTTACTTTGGTACTTTCAAT	250
Qy	101	AlAargLeuPheLeuGluAsnLeuilelySTyGlyyleuValAspProIleGlnVal	120
Db	251	GGACGGCGGCACTTGAAATATGTATCAAAATATGTATTTGATTAACACCCCTTCAGTGG	310
Qy	121	ValSerLeuPheLeuLlyAspProHis-----SerTTPProAlaProCysleuVal	137
Db	311	ATCTCAACGTTTGTGGACATCACTACATTCATTTGGAGCTGGCAAAATCTTGCTCTCATC	370
Qy	138	IleAlaIaIa 140	
Db	371	CTATGCTCA 379	
RESULT 9			
LOCUS	AR227755	381 bp	DNA
DEFINITION	Sequence 20 from patent US 6444876.	linear	PAT 20-DEC-2002
ACCESSION	AR227755		
VERSION	AR227755.1	GI:27266346	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 381)		
AUTHORS	Laesher,M.W. and Ruezinsky,D.M.		
TITLE	Acyl CoA: cholesterol acyltransferase related nucleic acid sequences		
JOURNAL	Patent: US 6444876-A 20 03-SBP-2002; Calgene LLC; Davis, CA; WOK;		
FEATURES			
source	Location/Qualifiers		
	1..381		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	4.11e-06	Length:	381
Score:	197.50	Matches:	41
Percent Similarity:	61.24	Conservative:	22
Best local Similarity:	39.84	Mismatches:	31
Query Match:	7.64	Indels:	9
DB:	6	Gaps:	2
US-10-659-800-6 (1-488) x AR227755 (1-381)			
Qy	47	ProAlaProAlaProAsn-----LysAspGlyAspAlaGlyValGly	60
Db	71	CCAGTAGCGCTCTCATTTGGCACAAATGTAATTCGCAAGGAAAGTTACAGAAATG	130
Qy	61	SeKglyHsITPgiuLeuArgCysHisArgLeuGlnAspSerLeuPheSerSerAspSer	80
Db	131	AGAGACCTTGCGAAGAGTGTATCATCTGCTCAAGATTATTTGTTTTCAGCAGTTCT	190
Qy	81	GlyPheSerAntyArxGlyyleuEunentPcySvalValMeleuileleuSerAn	100
Db	191	GGATGCAAAATTTCCGTGGATCTTCCTCAATTTGCTATTTTACTTTGGTACTTTCAAT	250
Qy	101	AlAargLeuPheLeuGluAsnLeuilelySTyGlyyleuValAspProIleGlnVal	120
Db	251	GGACGGCGGCACTTGAAATATGTATCAAAATATGTATTTGATTAACACCCCTTCAGTGG	310
Qy	121	ValSerLeuPheLeuLlyAspProHis-----SerTTPProAlaProCysleuVal	137
Db	311	ATCTCAACGTTTGTGGACATCACTACATTCATTTGGAGCTGGCAAAATCTTGCTCTCATC	370
Qy	138	IleAlaIaIa 140	
Db	371	CTATGCTCA 379	

[illegible]

MOX:

FEATURES

Location/Qualifiers

1..113

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 3,576-05 Length: 113

Score: 176.00 Matches: 36

Percent Similarity: 97.3% Conservative: 0

Best Local Similarity: 97.3% Mismatches: 1

Query Match: 6.8% Indels: 1

DB: 6 Gaps: 0

US-10-659-800-6 (1-488) x AR422734 (1-113)

QY 417 TTTTCTGATGAGGCTCCCTCTGCGAATGTTCCGCTCTGGGG-TTCACGGGCGATGATGGCT 61

DB 3 TACCTGATGAGGCTCCCTCTGCGAATGTTCCGCTCTGGGG-TTCACGGGCGATGATGGCT 61

QY 437 GlnIleProLeuAlaTrpPheValGlyArgPhePheGlnGlyAsnTyrGly 453

DB 62 CAGATCCACCTGGCTGTTCTGCGGCGCTTTTTCAGGGCACTATGGC 112

RESULT 12

AX983428 113 bp DNA linear PAT 15-JAN-2004

LOCUS AX983428 Sequence 14231 from Patent EP1104808.

DEFINITION AX983428

ACCESSION AX983428

VERSION AX983428.1 GI:40989568

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Dunas Milne Edwards, J.B., Jobert, S. and Giordano, J.Y.

AUTHORS ESTs and encoded human proteins

TITLE Patent: EP 1104808-A 14231 06-JUN-2001;

JOURNAL Genet (FR)

FEATURES

Location/Qualifiers

1..113

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 3,576-05 Length: 113

Score: 176.00 Matches: 36

Percent Similarity: 97.3% Conservative: 0

Best Local Similarity: 97.3% Mismatches: 1

Query Match: 6.8% Indels: 1

DB: 6 Gaps: 0

US-10-659-800-6 (1-488) x AX983428 (1-113)

QY 417 TTTTCTGATGAGGCTCCCTCTGCGAATGTTCCGCTCTGGGG-TTCACGGGCGATGATGGCT 61

DB 3 TACCTGATGAGGCTCCCTCTGCGAATGTTCCGCTCTGGGG-TTCACGGGCGATGATGGCT 61

QY 437 GlnIleProLeuAlaTrpPheValGlyArgPhePheGlnGlyAsnTyrGly 453

DB 62 CAGATCCACCTGGCTGTTCTGCGGCGCTTTTTCAGGGCACTATGGC 112

RESULT 13

BD241847 275 bp DNA linear PAT 17-JUL-2003

LOCUS BD241847

DEFINITION Acyl CoA:cholesterol acyltransferase related nucleic acid sequences.

ACCESSION BD241847

VERSION BD241847.1 GI:33051617

KEYWORDS JP 2002517201-A/4.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 275)

AUTHORS Laessner, M.W. and Ruezinsky, D.M.

TITLE Acyl CoA:cholesterol acyltransferase related nucleic acid sequences

JOURNAL Patent: JP 2002517201-A 4 18-JUN-2002;

COMMENT CALGENE LLC

OS Glycine max (soybean)

PN JP 2002517201-A/4

PD 18-JUN-2002

PF 04-JUN-1999 JP 2000552290

PR 05-JUN-1998 US 60/088143, 12-NOV-1998 US 60/108389 PI

MICHAEL W LAESSNER, DIANE M RUEZINSKY

PC C12N15/09, A01H5/00, C11B1/00, C11C3/00, C12N1/21, C12N5/10, C12N9/10, C12N15/00,

PC C12N5/00

CC n at positions 192, 202, 204, 211, 222, 234, 238, 239, 244,

CC 245, 247, 251,

CC 262, and 263 is unknown

CC key Location/Qualifiers

FT misc feature (192)..(263).

FEATURES

Location/Qualifiers

1..275

/organism="Glycine max"

/mol\_type="genomic DNA"

/db\_xref="taxon:3847"

ORIGIN

Alignment Scores:

Pred. No.: 0.000315 Length: 275

Score: 170.00 Matches: 34

Percent Similarity: 62.7% Conservative: 13

Best Local Similarity: 45.3% Mismatches: 28

Query Match: 6.6% Indels: 1

DB: 6 Gaps: 0

US-10-659-800-6 (1-488) x BD241847 (1-275)

QY 365 TTPAensergIuserValThrTrpGlnAntTrpAanIleProValHisLeuTrp 384

DB 2 TGGATGCCAAACCTGTTGAAGATTATTCAGAGATGGAATATGCCCTTCACAAATGG 61

QY 385 CysIleArgHisPheTyrLysProMetLeuArgGlySerSerLysTrpMetAlaArg 404

DB 62 ATGATCCGCCACCTATTATTTCCATGTTTAAGCACGGTATACCAAGCCGCTGCTT 121

QY 405 ThrGlyValPheLeuAlaSerValPhePheHisGluTyrLeuValSerValProLeuArg 424

DB 122 TTAATTCCTTCCTGCTGTTCTGCTTATTTCCATGAGCTGTCATCGCTGCTTGGCCA 180

QY 425 MetPheArgLeuTrpAlaPheThrGlyMetMetAlaGlnIlePro 439

DB 181 CATATTCAAGTNGTGGGTTTCNGNGCAATTTAAGTTTCAGTTCCT 225

RESULT 14

AR227742 275 bp DNA linear PAT 20-DEC-2002

LOCUS AR227742

DEFINITION Sequence 5 from patent US 644876.

ACCESSION AR227742

VERSION AR227742.1 GI:27266333

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 275)

AUTHORS Laessner, M.W. and Ruezinsky, D.M.

TITLE Acyl CoA:cholesterol acyltransferase related nucleic acid



sequences  
JOURNAL Patent: US 6444876-A 5 03-SEP-2002;  
Calgene LLC; Davis, CA;

W01;

FEATURES  
source Location/Qualifiers  
1..275  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 0.000315 Length: 275

Score: 170.00 Matches: 34

Percent Similarity: 62.7% Conservative: 13

Best Local Similarity: 45.3% Mismatches: 28

Query Match: 6.6% Indels: 1

DB: 6 Gaps: 0

US-10-659-800-6 (1-488) x AR227742 (1-275)

QY 365 TTPAenseGluSerValThrTyrPheTrpGlnAenTTPAenIleProValHisLysTrp 384

DB 2 TGGAAATGCCAAACCTGTGAAGATTATGAGAGATGTGAATATGCTTTCACAAATGG 61

QY 385 CysIleArgHisPheTyrLysProMetLeuArgGlySerSerLysTrpMetAlaArg 404

DB 62 ATGATCCCGCACCTATATTTTCCATGTTTAAGCAGCATATACCAAGCCGCTGCTCTT 121

QY 405 ThrGlyValPheLeuAlaSerAlaPhePheHisGlyTyrLeuValSerValProLeuArg 424

DB 122 TTAATTGCTTCCTGCTGT-TCGCTTTATTCATGAGCTGTGCATGCTGCTGCTGCCA 180

QY 425 MetPheArgLeuTrpAlaPheThrGlyMetMetAlaGlnIlePro 439

DB 181 CATATTCAAGTNGTGGGTTTCNGNGAATTNAGTTTCAGTNCCT 225

RESULT 15

AX090352

LOCUS Sequence 45 from Patent W00116308. 275 bp DNA linear PAT 21-MAR-2001

DEFINITION AX090352

ACCESSION AX090352.1 GI:13444213

KEYWORDS

SOURCE Glycine max (soybean)

ORGANISM

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

REFERENCE 1 Laessle, M. and van Eenennaam, A.

TITLE Plant sterol acyltransferases

JOURNAL Patent: WO 0116308-A 45 08-MAR-2001;

MONSANTO COMPANY (US)

FEATURES Location/Qualifiers

source 1..275

/organism="Glycine max"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3847"

1..275

/note="n=unknown"

ORIGIN

Alignment Scores:

Pred. No.: 0.000315 Length: 275

Score: 170.00 Matches: 34

Percent Similarity: 62.7% Conservative: 13

Best Local Similarity: 45.3% Mismatches: 28

Query Match: 6.6% Indels: 1

DB: 6 Gaps: 0

Search completed: May 7, 2006, 04:54:30

Job time : 7604 secs

QY 365 TTPAenseGluSerValThrTyrPheTrpGlnAenTTPAenIleProValHisLysTrp 384

DB 2 TGGAAATGCCAAACCTGTGAAGATTATGAGAGATGTGAATATGCTTTCACAAATGG 61

QY 385 CysIleArgHisPheTyrLysProMetLeuArgGlySerSerLysTrpMetAlaArg 404

DB 62 ATGATCCCGCACCTATATTTTCCATGTTTAAGCAGCATATACCAAGCCGCTGCTCTT 121

QY 405 ThrGlyValPheLeuAlaSerAlaPhePheHisGlyTyrLeuValSerValProLeuArg 424

DB 122 TTAATTGCTTCCTGCTGT-TCGCTTTATTCATGAGCTGTGCATGCTGCTGCTGCCA 180

QY 425 MetPheArgLeuTrpAlaPheThrGlyMetMetAlaGlnIlePro 439

DB 181 CATATTCAAGTNGTGGGTTTCNGNGAATTNAGTTTCAGTNCCT 225

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 7, 2006, 03:01:30 ; Search time 650 Seconds  
(without alignments)  
4586.303 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MGDGSSRRRTGSRPSSHG.....VLMYHYDYVLNYPAAEA 488

Scoring table: BLOSUM62  
Kgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 14491570

Minimum DB seq length: 0  
Maximum DB seq length: 500

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DBV=xlh  
-MODEL=frame.p2n.model -DBV=xlh  
-O=/dbs/ABSSWEB.spool/US10659800/runat\_05052006\_122320\_17509/app\_query.fasta\_1  
-DB=Published Applications NA\_New -QFMT=faetap -SUFFIX=p2n.szlm500.rnpbn  
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bites -START=1 -END=1  
-MATRIX=blomum62 -TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=0 -ALLEN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext  
-HEASIZE=500 -MINLEN=0 -MAXLEN=500 -HOST=absg05h  
-USER=US10659800@cgc1.1.660@runat\_05052006\_122320\_17509 -NCPU=6 -ICPU=3  
-NO\_WMAP -NGG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA\_New.\*

1: /SIDS5/ptodata/1/pubpna/US08\_NEW\_PUB.seq1.\*  
2: /SIDS5/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /SIDS5/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /SIDS5/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
5: /SIDS5/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /SIDS5/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
7: /SIDS5/ptodata/1/pubpna/US09\_NEW\_PUB.seq1.\*  
8: /SIDS5/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
9: /SIDS5/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
10: /SIDS5/ptodata/1/pubpna/US10\_NEW\_PUB.seq1.\*  
11: /SIDS5/ptodata/1/pubpna/US10\_NEW\_PUB.seq2.\*  
12: /SIDS5/ptodata/1/pubpna/US10\_NEW\_PUB.seq3.\*  
13: /SIDS5/ptodata/1/pubpna/US10\_NEW\_PUB.seq4.\*  
14: /SIDS5/ptodata/1/pubpna/US11\_NEW\_PUB.seq2.\*  
15: /SIDS5/ptodata/1/pubpna/US11\_NEW\_PUB.seq3.\*  
16: /SIDS5/ptodata/1/pubpna/US11\_NEW\_PUB.seq4.\*  
17: /SIDS5/ptodata/1/pubpna/US11\_NEW\_PUB.seq5.\*  
18: /SIDS5/ptodata/1/pubpna/US11\_NEW\_PUB.seq6.\*  
19: /SIDS5/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	166	6.4	201	10	US-10-995-561-13119	Sequence 13119, A
2	154	5.9	201	10	US-10-995-561-13118	Sequence 13118, A
3	149	5.7	201	10	US-10-995-561-13111	Sequence 13111, A
4	98	3.8	201	10	US-10-995-561-64752	Sequence 64752, A
5	96	3.7	201	10	US-10-995-561-64751	Sequence 64751, A
6	86	3.3	201	10	US-10-995-561-64750	Sequence 64750, A
7	83.5	3.2	444	13	US-10-983-809-47	Sequence 47, Appl
8	82.5	3.2	462	18	US-11-096-568A-12018	Sequence 12018, A
9	81	3.1	444	7	US-09-925-065A-159095	Sequence 159095,
10	81	3.1	452	12	US-10-301-480-252022	Sequence 252022,
11	81	3.1	452	12	US-10-301-480-865431	Sequence 865431,
12	81	3.1	469	7	US-09-925-065A-190003	Sequence 490003,
13	79	3.0	444	7	US-09-925-065A-159096	Sequence 159096,
14	79	3.0	452	12	US-10-301-480-252023	Sequence 252023,
15	79	3.0	452	12	US-10-301-480-865432	Sequence 865432,
16	79	3.0	468	7	US-09-925-065A-155469	Sequence 555469,
17	79	3.0	468	7	US-09-925-065A-555470	Sequence 555470,
18	78	3.0	401	17	US-11-000-688-850	Sequence 850, App
19	78	3.0	442	12	US-10-301-480-306219	Sequence 306219,
20	78	3.0	442	12	US-10-301-480-321978	Sequence 321978,
21	78	3.0	442	12	US-10-301-480-819628	Sequence 819628,
22	78	3.0	442	12	US-10-301-480-935387	Sequence 935387,
23	78	3.0	486	10	US-10-518-955-3	Sequence 3, Appl1
24	77.5	3.0	436	11	US-10-301-480-51306	Sequence 51306, A
25	77.5	3.0	436	11	US-10-301-480-51307	Sequence 51307, A
26	77.5	3.0	436	12	US-10-301-480-664715	Sequence 664715,
27	77.5	3.0	436	12	US-10-301-480-664716	Sequence 664716,
28	77.5	3.0	484	12	US-10-301-480-266676	Sequence 266676,
29	77.5	3.0	484	12	US-10-301-480-880085	Sequence 880085,
30	77.5	3.0	489	7	US-09-925-065A-175339	Sequence 175339,
31	77.5	3.0	493	7	US-09-925-065A-137506	Sequence 137506,
32	77.5	3.0	493	7	US-09-925-065A-137507	Sequence 137507,
33	77.5	3.0	493	7	US-09-925-065A-137508	Sequence 137508,
34	77.5	3.0	494	7	US-09-925-065A-923666	Sequence 923666,
35	77	3.0	487	17	US-11-108-172-140	Sequence 140, App
36	76.5	2.9	378	7	US-09-925-065A-86862	Sequence 86862, A
37	76.5	2.9	378	11	US-10-301-480-188102	Sequence 188102,
38	76.5	2.9	387	12	US-10-301-480-801511	Sequence 801511,
39	76.5	2.9	387	13	US-10-496-351-44	Sequence 44, Appl
40	76	2.9	500	18	US-11-036-196-1352	Sequence 1352, Ap
41	75.5	2.9	319	7	US-09-925-065A-48873	Sequence 48873, A
42	75.5	2.9	319	11	US-10-301-480-150111	Sequence 150111,
43	75.5	2.9	319	12	US-10-301-480-763520	Sequence 763520,
44	74.5	2.9	196	10	US-10-995-561-6062	Sequence 6062, Ap
45	74.5	2.9	249	17	US-11-043-752-3737	Sequence 3737, Ap

## ALIGNMENTS

RESULT 1  
US-10-995-561-13119  
; Sequence 13119, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13119  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-995-561-13119  
Alignment Scores: 6.77e-06 Length: 201  
Pred. No.: 1

Score: 166.00 Matches: 30  
Percent Similarity: 63.6% Conservative: 12  
Best Local Similarity: 45.5% Mismatches: 22  
Query Match: 6.4% Indels: 2  
DB: 10 Gaps: 1

US-10-659-800-6 (1-488) x US-10-995-561-13119 (1-201)

```
OY 353 MetGlnPheGlyAspArgGluPheTyrArgAspTrpTrpAsnSerGluSerValThrTyr 372
DB 3 CTACGATTTGGAGACAGATGTTCTACCGGAGCTGGTGAAGCTCAACGCTCTTCCAC 62
OY 373 PheTrpGlnAsnTrpAsnIleProValHisLysTrpCy1IleArgHisPheTyrLysPro 392
DB 63 TACTACCGGACTTGAACCTGGTGTGCTCATGACTGGCTTACAGCTACGTATACAGAT 122
OY 393 MetLeuArg-----ArgGlySerSerLysTrpMetAlaArgThrGlyValPheLeuAla 410
DB 123 GGGGTGGGGCTCTGTGTCGCCGGGCCGAGGGGTAGCATGCTGGGTGTGTTCTGTGTC 182
OY 411 SerAlaPhePheHisGlu 416
DB 183 TCCGCGATGGCCCATGAG 200
```

## RESULT 2

US-10-995-561-13118 Application US/10995561  
Publication No. US20050272054A1

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
FILE REFERENCE: CL001559  
CURRENT APPLICATION NUMBER: US/10/995,561  
CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13118  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-995-561-13118

## Alignment Scores:

Pred. No.: 8,17e-05 Length: 201  
Score: 154.00 Matches: 25  
Percent Similarity: 65.0% Conservative: 14  
Best Local Similarity: 41.7% Mismatches: 21  
Query Match: 5.9% Indels: 0  
DB: 10 Gaps: 0

US-10-659-800-6 (1-488) x US-10-995-561-13118 (1-201)

```
OY 319 IleIleGluArgLeuLeuValProAsnHisLeuIleTrpLeuIlePhePhe 338
DB 22 CTGGGCTCTATCTCGATGCCAGCTGCCAGGATCTTCATGCTGCTCATCTTC 81
OY 339 TyrTrpLeuPheHisSerCysLeuAsnAlaValAlaGluLeuMetGlnPheGlyAspArg 358
DB 82 TTGGCTTCTCTCATGCTGCTCAACGCTTTGGCGAGATGCTACGATTTGGAGACAG 141
OY 359 GluPheTyrArgAspTrpTrpAsnSerGluSerValThrTyrPheTrpGlnAsnTrpAsn 378
DB 142 ATGTCTACCGGAGCTGGAGACTCAACGCTCTTCTCCAACTACCGACCTTGGAAC 201
```

## RESULT 3

US-10-995-561-13111 Application US/10995561  
Publication No. US20050272054A1

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

FILE REFERENCE: CL001559

CURRENT APPLICATION NUMBER: US/10/995,561

CURRENT FILING DATE: 2004-11-24

NUMBER OF SEQ ID NOS: 85702

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13111

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-10-995-561-13111

## Alignment Scores:

Pred. No.: 0.000231 Length: 201  
Score: 149.00 Matches: 27  
Percent Similarity: 60.7% Conservative: 10  
Best Local Similarity: 44.3% Mismatches: 22  
Query Match: 5.7% Indels: 2  
DB: 10 Gaps: 1

US-10-659-800-6 (1-488) x US-10-995-561-13111 (1-201)

```
OY 360 PheTyrArgAspTrpTrpAsnSerGluSerValThrTyrPheTrpGlnAsnTrpAsnIle 379
DB 6 TTCTACCGGAGCTGGTGAAGCTCAACGCTCTTCCAACTACCTACCGGACTTGGAACTG 65
OY 380 ProValHisLysTrpCy1IleArgHisPheTyrLysProMetLeuArg-----ArgGly 397
DB 66 GTGCTCATGACTGGCTGTACAGCTCATGATACAGATGGAGGCTGGGCTCTTGGTCC 125
OY 398 SerSerLysTrpMetAlaArgThrGlyValPheLeuAlaSerAlaPhePheHisGluTyr 417
DB 126 CGGGCCGAGGAGGTATGACATGCTGGGTGTCTTCGTGCTCCGAGGTGCCCATGAGATAT 185
OY 418 Leu 418
DB 186 ATC 188
```

## RESULT 4

US-10-995-561-64752 Application US/10995561  
Publication No. US20050272054A1

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

FILE REFERENCE: CL001559

CURRENT APPLICATION NUMBER: US/10/995,561

CURRENT FILING DATE: 2004-11-24

NUMBER OF SEQ ID NOS: 85702

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 64752

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-10-995-561-64752

## Alignment Scores:

Pred. No.: 9.11 Length: 201  
Score: 98.00 Matches: 16  
Percent Similarity: 56.8% Conservative: 9  
Best Local Similarity: 36.4% Mismatches: 15  
Query Match: 3.8% Indels: 4  
DB: 10 Gaps: 1

US-10-659-800-6 (1-488) x US-10-995-561-64752 (1-201)

```
OY 362 ArgAspTrpTrpAsnSerGluSerValThrTyrPheTrpGlnAsnTrpAsnIleProVal 381
DB 30 CAGGACTGTGTGAAGCTCAACGCTCTTCTCCAACTACCGACCTTGGAACTGTGTGTC 89
```

```

Qy 382 HlslvSTPcYsllelrgHshPheTyrlYsPromeleuAargAlgYserserlySTP 401
Db 90 CAGACTGGCTGTACAGCTACGTGTATCAGGATGGCTGCGG-----GATGG 137
Qy 402 MetAlaArgThr 405
Db 138 GCCCTGCAGACC 149

RESULT 5
US-10-995-561-64711
; Sequence 64711: Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ. ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64711
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-64711

Alignment Scores:
Pred. No.: 13.8 length: 201
Score: 96.00 Matches: 16
Percent Similarity: 54.5% Conservative: 8
Best Local Similarity: 36.4% Mismatches: 16
Query Match: 3.7% Indels: 4
DB: 10 Gaps: 1

US-10-659-800-6 (1-488) x US-10-995-561-64711 (1-201)
Qy 362 ArgAspTrpTrpAsnSerGluSerValThrTyPheTrpGlnAsnTrpAsnIleProval 381
Db 12 CAGGACTGGTGGAACTCAACGTCCTTCGCACTACTACCGGCACTTGGAACTGGTGGTC 71
Qy 382 HlslvSTPcYsllelrgHshPheTyrlYsPromeleuAargAlgYserserlySTP 401
Db 72 CATGACTGGCTGTACAGCTACGTGTATCAGGATGGCTGCGG-----GATGG 119
Qy 402 MetAlaArgThr 405
Db 120 GCCCTGCAGACC 131

RESULT 6
US-10-995-561-64750
; Sequence 64750: Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ. ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64750
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-64750

Alignment Scores:
Pred. No.: 11.0 length: 201
Score: 86.00 Matches: 16

```

```

Percent Similarity: 64.7%
Best Local Similarity: 47.1%
Query Match: 3.3%
DB: 10

US-10-659-800-6 (1-488) x US-10-995-561-64750 (1-201)

QY 329 ProAhsHsleuIeTtRpleHlePhnePhetYtTtRpleuhenHsSerCySleuAenA 348
DB 52 CCAGGCACTTCACGCTGCCTCATCTTCTTCCTTCCTCATCTTCACGCTTCACGCC 111
QY 349 ValAaIeureMetGInPheGlyAspArgGluPheTyrArg 362
DB 112 TTTCGAGAGTCACTACGATTGGAGACAGATGTTCTACCG 153

RESULT 7
US-10-983-809-47
; Sequence 47, Application US/10983809
; Publication No. US20060063164A1
; GENERAL INFORMATION:
; APPLICANT: An, Sungwhan
; APPLICANT: Yoon, ChiWang
; APPLICANT: Oh, Taejeong
; APPLICANT: Yoon, Daekyoung
; APPLICANT: Lee, Sunwoo
; APPLICANT: Kim, Myungsoon
; APPLICANT: Woo, Sukkyung
; TITLE OF INVENTION: METHOD FOR DETECTING METHYLATION OF PROMOTER USING RESTRICTION
; FILE REFERENCE: 4240-113
; CURRENT APPLICATION NUMBER: US/10/983, 809
; CURRENT FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: KR 10-2004-0075395
; PRIOR FILING DATE: 2004-09-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-983-809-47

Alignment Scores:
Pred. No.: 465 Length: 444
Score: 83.50 Matches: 28
Percent Similarity: 47.2% Conservative: 6
Best Local Similarity: 38.9% Mismatches: 29
Query Match: 3.2% Indels: 9
DB: 13 Gaps: 4

US-10-659-800-6 (1-488) x US-10-983-809-47 (1-444)

QY 2 G1yAspArgGlySerSerArg---ArgATgArgThGlySerArgProSerSerHigly 20
DB 141 GGGAGGTGCGGGGGCGGAGGAGCGAAGGAGGAGGAGGAGG---ACCGGAGGGGT 197
QY 21 G1yGlyGlyProAlaAlaAlaIaGluGluGluValArgAspAlaAlaIaGlyProAspVal 40
DB 198 GCTGGCGGGGGGTGCTGTGTGGGTGAGAAAGCCGCTTAGCAAAATTTGGGGCCGACAG 257
QY 41 G1yAlaAlaGlyAspAlaProAlaPro-----AlaProAsnlyAspGly--- 55
DB 258 GCAGCACTCGGCTTTAACTTGAGCACTGAAGCGGGGGAAGAAAGCAAAAGAAAGGGGTG 317
QY 56 -----AspAlaGlyValAlaGlySerGlyHsITPglu 65
DB 318 GTGTGCGAGTAgGGGTGGGTGGGGGGAATTTGAG 353

RESULT 8
US-11-096-568A-12018/C
; Sequence 12018, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thekey
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12018
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(462)
; OTHER INFORMATION: Ceres Seq. ID no. 13660322
US-11-096-568A-12018

Alignment Scores:
Pred. No.: 600 Length: 462
Score: 82.50 Matches: 36
Percent Similarity: 38.3% Conservative: 8
Best Local Similarity: 31.3% Mismatches: 54
Query Match: 3.2% Indels: 17
DB: 18 Gaps: 4

US-10-659-800-6 (1-488) x US-11-096-568A-12018 (1-462)
QY 2 G1YAPAG1G1YSerSerArgArgG1YThrG1YSerArgProSerSerHiSg1Y 21
DB 324 GGGGACGGGCGCGAAGTGGCGGCGCGCGCGCGCGCGCGCTGTGTGTGCG 265
QY 22 ---G1YGLYProAlaAlaGluGluValArgAspAlaAlaAlaG1YProAspAl 40
DB 264 CGGAGTGAAGCAGCCGCTTCCAGCCGCCCGCGCGCGAGCCCGAGCTGACACTCA 205
QY 41 G1YAlaAlaG1YAspAlaPro-----AlaProAlaProAsnLysAspG1YAspAla 57
DB 204 GCAGCGGCGAGCCCGCAGCGCAGCATGCGCGCGCGGGAATTCGACGAGCAGGTC 145
QY 58 G1YValG1YSerG1YHiStrG1YLeuArgCySHiStrG1YLeuGlnAspSerLeuPhe 77
DB 144 TCGGGCGGCGCGCGCGCAGAGAGATTCGCGCGGTGCGCATCGATCCGCGGG 85
QY 78 SerAspSerG1YPheSerAsnTyraG1YIleLeuAsnTrpCyValValMetLeuIle 97
DB 84 ACTAGCTGGGAGCAGAGAGATTTGGCGGATTTA-----TTAATT 43
QY 98 LeuSerAsnAlaArgLeuPheLeuGluAsnLeuIleLysTyG1Y 112
DB 42 CGGAGTGAAGTGAAGCA-----AAATACGGA 19

RESULT 9
US-09-925-065A-159095
; Sequence 159095, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159095
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-159095

Alignment Scores:
Pred. No.: 782 Length: 444
Score: 81.00 Matches: 33
Percent Similarity: 36.5% Conservative: 2
Best Local Similarity: 34.4% Mismatches: 49
Query Match: 3.1% Indels: 12
DB: 7 Gaps: 4

US-10-659-800-6 (1-488) x US-09-925-065A-159095 (1-444)
QY 8 ArgArgArgArgThrG1YSerArgProSer-----SerHiSg1YG1YG1YPro--- 24
DB 124 AGACGAGCAATCAAGCGCGCGCGCGCGCGCTGCTGATGACGGGCAAGCGGCCCTTT 183
QY 25 -----AlaAlaAlaGluGluValArgAspAlaAlaAlaG1YProAspValG1YAla 42
DB 184 CCCGACCGCGCGCAGAGGAGAGAAACCGGCGCTCGCCCGAGCCTTCCAGAGCAGCG 243
QY 43 AlaG1YAspAlaProAlaProAlaProAsnLysAspG1YAspAlaG1YValG1YSerG1Y 62
DB 244 GCGCGGAGAGAGAGCGCGCGCGCGCGCGGAGCAACGCGCGCGCGCAGAGCTCGGG 303
QY 63 -----HiStrG1YLeuArgCySHiStrG1YLeuGlnAspSerLeuPhe 76
DB 304 CCAGCGGCGCGCATCCCGCGCGCGCTGCGCGGTGAGAGCTTGCAGGCGCGCGGTG 363
QY 77 SerSerAsp---SerG1YPheSerAsnTyraG1YIleLeuAsnTrp 91
DB 364 CCGCAGCAGAGAGTGAACAAGATGCGAGGATCTTAAGCTGG 411

RESULT 10
US-10-301-480-252022
; Sequence 252022, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1225818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252022
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-301-480-252022

Alignment Scores:
Pred. No.: 798 Length: 452
Score: 81.00 Matches: 33
Percent Similarity: 36.5% Conservative: 2
Best Local Similarity: 34.4% Mismatches: 49
Query Match: 3.1% Indels: 12
DB: 12 Gaps: 4

US-10-659-800-6 (1-488) x US-10-301-480-252022 (1-452)
QY 8 ArgArgArgArgThrG1YSerArgProSer-----SerHiSg1YG1YG1YPro--- 24
DB 124 AGACGAGCAATCAAGCGCGCGCGCGCGCTGCTGATGACGGGCAAGCGGCCCTTT 183
```

```
Qy 25 -----AlaAlaAlaGluGluValAlaArgAspAlaAlaAlaGlyProAspValAlaAla 42
Db 164 CCCACCGCGCGCCAGAGGAGAGAGAACCCGGGCTCCGCCGAGCTTCGAGAGACGG 243
Qy 43 AlAGlyAspAlaProAlaProAlaProAlaProAlaProAlaGlyAspAlaGlyValGlySergly 62
Db 244 GCCCGGAGAGAGCGCGCGCGCGCGGAGACGAGCGCGCGCGCGCACAGGCTCGGG 303
Qy 63 -----HistrgLuleuArgCySHIAArgLeuGlnAspSerLeuPhe 76
Db 304 CCAGCGCGCGCGCATCCCGCGCGCGCTGCGGTGAGAGACTTGGCGGCTGCGGGTG 363
Qy 77 SerSerAsp---SergLyPheSerAsnTrpArgGlyIleLeuAsnTrp 91
Db 364 CCCGAGACAGAGAGTGGACAAAGCAAGATGGCAGGATTTAGCTGG 411

RESULT 11
US-10-301-480-865431
; Sequence 865431, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIORITY FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 126818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 865431
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-865431

Alignment Scores:
Pred. No.: 798 Length: 452
Score: 81.00 Matches: 33
Percent Similarity: 36.5% Conservative: 2
Best Local Similarity: 34.4% Mismatches: 49
Query Match: 3.1% Indels: 12
Gaps: 4
DB: 12

US-10-659-800-6 (1-488) x US-10-301-480-865431 (1-452)
Qy 8 ArgArgArgArgThgThgSeraArgProSer-----SerHisGlyGlyGlyPro--- 24
Db 124 AGACGAGCAATCAAGGCGCGCGCGCGCGAGCTGCATGTGACGGGCAAGCGGCCCTTT 183
Qy 25 -----AlaAlaAlaGluGluValAlaArgAspAlaAlaAlaGlyProAspValAlaAla 42
Db 184 CCCACCGCGCGCCAGAGGAGAGAGAACCCGGGCTCCGCCGAGCTTCGAGAGACGG 243
Qy 43 AlAGlyAspAlaProAlaProAlaProAlaProAlaProAlaGlyAspAlaGlyValGlySergly 62
Db 244 GCCCGGAGAGAGCGCGCGCGCGCGGAGACGAGCGCGCGCGCGCACAGGCTCGGG 303
Qy 63 -----HistrgLuleuArgCySHIAArgLeuGlnAspSerLeuPhe 76
Db 304 CCAGCGCGCGCGCATCCCGCGCGCGCTGCGGTGAGAGACTTGGCGGCTGCGGGTG 363
Qy 77 SerSerAsp---SergLyPheSerAsnTrpArgGlyIleLeuAsnTrp 91
Db 364 CCCGAGACAGAGAGTGGACAAAGCAAGATGGCAGGATTTAGCTGG 411

RESULT 12
US-09-925-065A-490003
; Sequence 490003, Application US/09925065A
```

```
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIORITY FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 490003
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-490003

Alignment Scores:
Pred. No.: 933 Length: 469
Score: 81.00 Matches: 30
Percent Similarity: 39.6% Conservative: 25
Best Local Similarity: 21.6% Mismatches: 49
Query Match: 3.1% Indels: 35
Gaps: 5
DB: 7

US-10-659-800-6 (1-488) x US-09-925-065A-490003 (1-469)
Qy 174 ThrIleLeuCyPheProAlaAlaValAlaLeuLeuValGluSerIleThrProValGly 193
Db 66 AGTGTCTTGTGCTTCTCTCC-----ACCGAATTTT 98
Qy 194 SerLeuLeuAlaLeuMetAlaHisThrIleLeuPheLeuValGluPheSerTrpArgAsp 213
Db 99 GCTCTTCTTCAACGAGCAATATGATGCTTATTCATGAATAATTTCTTAATGCTTAA 158
Qy 214 ValAsn-----SerTrpCyArgArgAlaArgAlaValAlaSerAlaGlyValys 231
Db 159 ATMAACTGCAGAGTGTGGTGGAGCACTGAACACATACCTTGAAGCTTATCTAATAG 218
Qy 232 AlaSerSerAlaAlaAlaProHisThrValSerTrpProAspAsnLeu---ThrTrpArg 250
Db 219 TTGAATCATAGTATATATATGTCATGTATCTCCCTCAGGTAGTATGATTCAG 278
Qy 251 AspLeuTrpTrpPheLeuPheAlaProThrLeuCySryGluLeuAsnPheProArgSer 270
Db 279 GACATAAATATTTTATTTAT-----TGTACCAAGTTTCTTT----- 317
Qy 271 ProArgIleAlaGlyArgPheLeuLeuArgArgIleLeuGluMetLeuPheThrGln 290
Db 318 -----ATACTATTCTCTAAT 332
Qy 291 LeuGlnValGlyLeuIle-GInGInTrpMetValProThrIleGlnAsnSerMet 308
Db 333 TTAGAATGTGATTTTTCACAGTATGGGTTCGCAAGAGTAAAGAACTAAATAG 387

RESULT 13
US-09-925-065A-159096
; Sequence 159096, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
```

CURRENT APPLICATION NUMBER: US/09/925,065A  
CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 159096  
LENGTH: 444  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-159096

Alignment Scores:  
Pred. No.: 1,18e+03 Length: 444  
Score: 79.00 Matches: 33  
Percent Similarity: 35.4% Conservative: 1  
Best Local Similarity: 34.4% Mismatches: 50  
Query Match: 3.0% Indels: 12  
Gaps: 4

US-10-659-800-6 (1-488) x US-09-925-065A-159096 (1-444)

QY 8 ArgATGATGATGThGlySerATGProSer-----SerHISGLYGLYGLYPro--- 24  
DB 124 AGACGAGCCAAATCAGGCGCGCGGCCGCCAGCTGAGCGGGAAGCGGCCCTTT 183  
QY 25 -----AAlaAlaAGluGluValAlaArgAspAlaAlaAlaGlyProAspValGlyAla 42  
DB 184 CCCGACCGCGCCAGAGGAGGAGAGAACCGGGGCTCCCGGAGCTTCGAGAGCAGCG 243  
QY 43 AlaGlyAspAlaProAlaProAlaProAlaProAlaAspGlyAspAlaGlyValGlySerGly 62  
DB 244 GCCGCGAGAGAGCGCGCGCGCGCGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 303  
QY 63 -----HISTrpGluLeuATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76  
DB 304 CCAGCGCGCGCGCGCATCCCGCGCGCGCTCGCGCGGAGAGCTTGGCGGCGTGGCGG 363  
QY 77 SerSerAsp---SerGlyPheSerAsnTrpArgGlyIleLeuAsnTrp 91  
DB 364 CCGCAGACAGAGATGAGCAAGCAAGATGCGAGGATCTTAGCCCTGG 411

RESULT 14

US-10-301-480-252023  
Sequence 252023, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-06-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 252023  
LENGTH: 452  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-252023

Alignment Scores:  
Pred. No.: 1,21e+03 Length: 452  
Score: 79.00 Matches: 33  
Percent Similarity: 35.4% Conservative: 1  
Best Local Similarity: 34.4% Mismatches: 50  
Query Match: 3.0% Indels: 12  
Gaps: 4

US-10-659-800-6 (1-488) x US-10-301-480-252023 (1-452)

QY 8 ArgATGATGATGThGlySerATGProSer-----SerHISGLYGLYGLYPro--- 24  
DB 124 AGACGAGCCAAATCAGGCGCGCGGCCGCCAGCTGAGCGGGAAGCGGCCCTTT 183  
QY 25 -----AAlaAlaAGluGluValAlaArgAspAlaAlaAlaGlyProAspValGlyAla 42  
DB 184 CCCGACCGCGCCAGAGGAGGAGAGAACCGGGGCTCCCGGAGCTTCGAGAGCAGCG 243  
QY 43 AlaGlyAspAlaProAlaProAlaProAlaProAlaAspGlyAspAlaGlyValGlySerGly 62  
DB 244 GCCGCGAGAGAGCGCGCGCGCGCGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 303  
QY 63 -----HISTrpGluLeuATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76  
DB 304 CCAGCGCGCGCGCATCCCGCGCGCGCTCGCGCGGAGAGCTTGGCGGCTGGCGG 363  
QY 77 SerSerAsp---SerGlyPheSerAsnTrpArgGlyIleLeuAsnTrp 91  
DB 364 CCGCAGACAGAGATGAGCAAGCAAGATGCGAGGATCTTAGCCCTGG 411

RESULT 15

US-10-301-480-865432  
Sequence 865432, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-06-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 865432  
LENGTH: 452  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-865432

Alignment Scores:  
Pred. No.: 1,21e+03 Length: 452  
Score: 79.00 Matches: 33  
Percent Similarity: 35.4% Conservative: 1  
Best Local Similarity: 34.4% Mismatches: 50  
Query Match: 3.0% Indels: 12  
Gaps: 4

US-10-659-800-6 (1-488) x US-10-301-480-865432 (1-452)

QY 8 ArgATGATGATGThGlySerATGProSer-----SerHISGLYGLYGLYPro--- 24  
DB 124 AGACGAGCCAAATCAGGCGCGCGGCCGCCAGCTGAGCGGGAAGCGGCCCTTT 183  
QY 25 -----AAlaAlaAGluGluValAlaArgAspAlaAlaAlaGlyProAspValGlyAla 42  
DB 184 CCCGACCGCGCCAGAGGAGGAGAGAACCGGGGCTCCCGGAGCTTCGAGAGCAGCG 243  
QY 43 AlaGlyAspAlaProAlaProAlaProAlaProAlaAspGlyAspAlaGlyValGlySerGly 62  
DB 184 CCCGACCGCGCCAGAGGAGGAGAGAACCGGGGCTCCCGGAGCTTCGAGAGCAGCG 243



Db 244 GCTCGGAGAGAGCGCGCGCGCGCGCGGAGACATCGCGCGCGCGACAGGCTTCGGG 303  
 GY 63 -----HistTropIuLeuAaGcYshIaArgLeuGlnAhpSerLeuPhe 76  
 Db 304 CCAAGCGGCGCGGCATCCCCGAGCGCGCCCTCGCGCGAGAGAGAGCTTGGCGGCGCTGCGGAGT 363  
 GY 77 SerSerAap--SerGlyPheSerAenTyzArgGlyY1LeuAaSnTP 91  
 Db 364 CCGCAGAGACAGGATGGACAAAGACAGATGCGCGGATCTTGAGCCCTGG 411

Search completed: May 7, 2006, 03:12:30  
Job time : 653 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2006, 12:44:20 ; Search time 55 Seconds  
(without alignments)  
3707.286 Million cell updates/sec

Title: US-10-659-800-6  
Perfect score: 2594  
Sequence: 1 MGDGSSRRRTGSRPSSHG.....VLMYVHYVLYNEAPAAEA 488

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBSCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBSCOMB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBSCOMB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBSCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBSCOMB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBSCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2594	100.0	488	US-10-273-438-6	Sequence 6, Appl1
2	2594	100.0	488	US-10-040-315A-6	Sequence 6, Appl1
3	2594	100.0	488	US-10-307-817-544	Sequence 544, App
4	2594	100.0	488	US-10-659-800-6	Sequence 6, Appl1
5	2594	100.0	488	US-11-117-005-5	Sequence 5, Appl1
6	2588	99.8	488	US-10-741-600-1339	Sequence 1339, Ap
7	2588	99.8	488	US-10-741-600-1341	Sequence 1341, Ap
8	2299.5	88.6	500	US-10-157-855-17	Sequence 17, Appl
9	2299.5	88.6	500	US-10-647-517-24	Sequence 24, Appl
10	2292.5	88.4	489	US-10-415-620-2	Sequence 2, Appl1
11	2292.5	88.4	489	US-10-415-620-5	Sequence 5, Appl1
12	2278.5	87.8	498	US-10-482-936-4	Sequence 4, Appl1
13	2278.5	87.8	498	US-11-117-005-4	Sequence 4, Appl1
14	2273.5	87.6	498	US-10-273-438-7	Sequence 7, Appl1
15	2273.5	87.6	498	US-10-273-438-10	Sequence 10, Appl1
16	2273.5	87.6	498	US-10-040-315A-7	Sequence 7, Appl1
17	2273.5	87.6	498	US-10-040-315A-10	Sequence 10, Appl1
18	2273.5	87.6	498	US-10-659-800-7	Sequence 7, Appl1
19	2273.5	87.6	498	US-10-659-800-10	Sequence 10, Appl1
20	2272.5	87.6	498	US-10-482-936-2	Sequence 2, Appl1
21	2264	87.3	497	US-10-690-994-25	Sequence 25, Appl1
22	2192.5	84.5	467	US-10-415-620-6	Sequence 6, Appl1
23	2046	78.9	387	US-10-741-600-1340	Sequence 1340, Ap
24	2039	78.6	386	US-10-273-438-5	Sequence 5, Appl1
25	2039	78.6	386	US-10-040-315A-5	Sequence 5, Appl1
26	2039	78.6	386	US-10-659-800-5	Sequence 5, Appl1
27	1681	64.8	421	US-10-741-600-1342	Sequence 1342, Ap

28	1297	50.0	236	3	US-09-764-853-499	Sequence 499, App
29	1106	42.6	496	4	US-10-157-855-18	Sequence 18, Appl
30	1106	42.6	496	4	US-10-369-493-6484	Sequence 6484, Ap
31	1106	42.6	496	4	US-10-369-493-6485	Sequence 6485, Ap
32	1106	42.6	496	5	US-10-647-517-14	Sequence 14, Appl
33	840	32.4	517	4	US-10-424-599-254338	Sequence 254338, A
34	834	32.2	538	4	US-10-425-114-49477	Sequence 49477, A
35	832.5	32.1	501	4	US-10-103-247-2	Sequence 2, Appl1
36	819	31.6	507	5	US-10-859-247-2	Sequence 2, Appl1
37	819	31.6	507	5	US-10-859-247-17	Sequence 17, Appl1
38	807	31.1	508	4	US-10-690-994-22	Sequence 22, Appl1
39	800.5	30.9	504	4	US-10-690-994-16	Sequence 16, Appl1
40	798.5	30.8	520	4	US-10-690-994-2	Sequence 2, Appl1
41	797	30.7	532	4	US-10-223-076-13	Sequence 13, Appl1
42	797	30.7	532	5	US-10-859-247-19	Sequence 19, Appl1
43	791.5	30.5	520	4	US-10-157-855-2	Sequence 2, Appl1
44	791.5	30.5	520	4	US-10-223-076-3	Sequence 3, Appl1
45	791.5	30.5	520	4	US-10-690-994-26	Sequence 26, Appl1

ALIGNMENTS

```
RESULT 1
US-10-273-438-6
; Sequence 6, Application US/10273438
; Publication No. US2003007257A1
; GENERAL INFORMATION:
; APPLICANT: Farese, Robert V.
; APPLICANT: Casee, Sylvaine
; APPLICANT: Smith, Steven
; APPLICANT: Erickson, Sandra
; TITLE OF INVENTION: Diacylglycerol O-Acyltransferase
; FILE REFERENCE: UCAL-105C1P2
; CURRENT FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US/10/040,315
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/107,771
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: PCT/US98/17883
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/103,754
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 09/339,472
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 488
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-273-438-6
Query Match      100.0%; Score 2594; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 8.9e-23;
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MGDGSSRRRTGSRPSSHGCGGPAABEVRDAAAGPDVGAAGDAPAPAPNKGDAVG 60
DB      1 MGDGSSRRRTGSRPSSHGCGGPAABEVRDAAAGPDVGAAGDAPAPAPNKGDAVG 60
QY      61 SGHWEIRCHRIODSLFSSDSGFSNYRGILMNCVWLTLSNARLFLENLITKGIIVDP1QV 120
DB      61 SGHWEIRCHRIODSLFSSDSGFSNYRGILMNCVWLTLSNARLFLENLITKGIIVDP1QV 120
QY      121 VSLFKDHPAPCLVIAAVFAVFAFQVEKRLAVGALTEOAGILLHVAHLATILCFPA 180
DB      121 VSLFKDHPAPCLVIAAVFAVFAFQVEKRLAVGALTEOAGILLHVAHLATILCFPA 180
QY      181 AVVILVESITPVGSILMAHFTIILFLKFSYRDVNSWCRARAPAAAGAKKASAAAPHT 240
DB      181 AVVILVESITPVGSILMAHFTIILFLKFSYRDVNSWCRARAPAAAGAKKASAAAPHT 240
```

QY 241 VSYPDNLTYRDLVYFLFAPPTLCYELNPPSPRIKRPFLRLRIEMLPFTOLQVGLIQOMM 300  
DB 241 VSYPDNLTYRDLVYFLFAPPTLCYELNPPSPRIKRPFLRLRIEMLPFTOLQVGLIQOMM 300  
QY 301 VPTI0NSMKPKPMDYSRIIERLLKLAVPNHLIWLIFPYWLFHSCINAVALMOFGDREF 360  
DB 301 VPTI0NSMKPKPMDYSRIIERLLKLAVPNHLIWLIFPYWLFHSCINAVALMOFGDREF 360  
QY 361 YRDWNNSESVTYFMQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFPHXYLV 420  
DB 361 YRDWNNSESVTYFMQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFPHXYLV 420  
QY 421 VPLRMFRLMFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQOPIAVLMYVHDYVYLN 480  
DB 421 VPLRMFRLMFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQOPIAVLMYVHDYVYLN 480  
QY 481 YEAPAAEA 488  
DB 481 YEAPAAEA 488

RESULT 2  
US-10-040-315A-6  
; Sequence 6, Application US/10040315A  
; Publication No. US20030167483A1  
; GENERAL INFORMATION:  
; APPLICANT: Fareese, Robert V.  
; APPLICANT: Caees, Sylvaine  
; APPLICANT: Smith, Steven  
; APPLICANT: Erickson, Sandra  
; TITLE OF INVENTION: Diacylglycerol O-Acyltransferase  
; FILE REFERENCE: UCAL-105CIP2  
; CURRENT APPLICATION NUMBER: US/10/040,315A  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/107,771  
; PRIOR FILING DATE: 1998-11-09  
; PRIOR APPLICATION NUMBER: PCT/US98/17883  
; PRIOR FILING DATE: 1998-08-28  
; PRIOR APPLICATION NUMBER: 09/103,754  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 09/339,472  
; PRIOR FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 488  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-040-315A-6

Query Match 100.0%; Score 2594; DB 4; Length 488;  
Best Local Similarity 100.0%; Pred. No. 8.9e-233;  
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRRRTGSRPSSHGCGGPAABEEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60  
DB 1 MGDGSSRRRRRTGSRPSSHGCGGPAABEEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60  
QY 61 SGHWEILRCHRLQDSLFSSDSGFSNYRGILNMCVWMLISNARLFLENLIKYGILVDP1QV 120  
DB 61 SGHWEILRCHRLQDSLFSSDSGFSNYRGILNMCVWMLISNARLFLENLIKYGILVDP1QV 120  
QY 121 VSLFKDPSHWPAPCLVIAANVFAVAAFOVEKRLAVGALTQAGLLHVANLATILCPFA 180  
DB 121 VSLFKDPSHWPAPCLVIAANVFAVAAFOVEKRLAVGALTQAGLLHVANLATILCPFA 180  
QY 181 AVLLVSEITVPGSLALMAHTILFLKLFYRDVNSWCRRARAKAASGKASAAAPHT 240  
DB 181 AVLLVSEITVPGSLALMAHTILFLKLFYRDVNSWCRRARAKAASGKASAAAPHT 240  
QY 241 VSYPDNLTYRDLVYFLFAPPTLCYELNPPSPRIKRPFLRLRIEMLPFTOLQVGLIQOMM 300  
DB 241 VSYPDNLTYRDLVYFLFAPPTLCYELNPPSPRIKRPFLRLRIEMLPFTOLQVGLIQOMM 300

DB 241 VSYPDNLTYRDLVYFLFAPPTLCYELNPPSPRIKRPFLRLRIEMLPFTOLQVGLIQOMM 300  
QY 301 VPTI0NSMKPKPMDYSRIIERLLKLAVPNHLIWLIFPYWLFHSCINAVALMOFGDREF 360  
DB 301 VPTI0NSMKPKPMDYSRIIERLLKLAVPNHLIWLIFPYWLFHSCINAVALMOFGDREF 360  
QY 361 YRDWNNSESVTYFMQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFPHXYLV 420  
DB 361 YRDWNNSESVTYFMQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFPHXYLV 420  
QY 421 VPLRMFRLMFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQOPIAVLMYVHDYVYLN 480  
DB 421 VPLRMFRLMFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQOPIAVLMYVHDYVYLN 480  
QY 481 YEAPAAEA 488  
DB 481 YEAPAAEA 488

RESULT 3  
US-10-307-817-544  
; Sequence 544, Application US/10307817  
; Publication No. US20040058338A1  
; GENERAL INFORMATION:  
; APPLICANT: Agree et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-502C  
; CURRENT APPLICATION NUMBER: US/10/307,817  
; CURRENT FILING DATE: 2002-12-02  
; NUMBER OF SEQ ID NOS: 682  
; SOFTWARE: Curaseqlist version 0.1  
; SEQ ID NO 544  
; LENGTH: 488  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-307-817-544

Query Match 100.0%; Score 2594; DB 4; Length 488;  
Best Local Similarity 100.0%; Pred. No. 8.9e-233;  
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRRRTGSRPSSHGCGGPAABEEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60  
DB 1 MGDGSSRRRRRTGSRPSSHGCGGPAABEEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60  
QY 61 SGHWEILRCHRLQDSLFSSDSGFSNYRGILNMCVWMLISNARLFLENLIKYGILVDP1QV 120  
DB 61 SGHWEILRCHRLQDSLFSSDSGFSNYRGILNMCVWMLISNARLFLENLIKYGILVDP1QV 120  
QY 121 VSLFKDPSHWPAPCLVIAANVFAVAAFOVEKRLAVGALTQAGLLHVANLATILCPFA 180  
DB 121 VSLFKDPSHWPAPCLVIAANVFAVAAFOVEKRLAVGALTQAGLLHVANLATILCPFA 180  
QY 181 AVLLVSEITVPGSLALMAHTILFLKLFYRDVNSWCRRARAKAASGKASAAAPHT 240  
DB 181 AVLLVSEITVPGSLALMAHTILFLKLFYRDVNSWCRRARAKAASGKASAAAPHT 240  
QY 241 VSYPDNLTYRDLVYFLFAPPTLCYELNPPSPRIKRPFLRLRIEMLPFTOLQVGLIQOMM 300  
DB 241 VSYPDNLTYRDLVYFLFAPPTLCYELNPPSPRIKRPFLRLRIEMLPFTOLQVGLIQOMM 300  
QY 301 VPTI0NSMKPKPMDYSRIIERLLKLAVPNHLIWLIFPYWLFHSCINAVALMOFGDREF 360  
DB 301 VPTI0NSMKPKPMDYSRIIERLLKLAVPNHLIWLIFPYWLFHSCINAVALMOFGDREF 360  
QY 361 YRDWNNSESVTYFMQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFPHXYLV 420  
DB 361 YRDWNNSESVTYFMQNNNIPVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFPHXYLV 420  
QY 421 VPLRMFRLMFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQOPIAVLMYVHDYVYLN 480  
DB 421 VPLRMFRLMFTGMAAOIPLAMFVGRFPGQNYGNAAVWLSLIIQOPIAVLMYVHDYVYLN 480

QY 481 YEAPAAEA 488  
Db 481 YEAPAAEA 488

RESULT 4  
US-10-659-800-6  
; Sequence 6, Application US/10659800  
; Publication No. US20040078836A1  
; GENERAL INFORMATION:  
; APPLICANT: Fatese, Robert V.  
; APPLICANT: Cases, Sylvaine  
; APPLICANT: Smith, Steven  
; APPLICANT: Erickson, Sandra  
; TITLE OF INVENTION: Diacylglycerol O-Acyltransferase  
; FILE REFERENCE: UCAL-105CIP2CON2  
; CURRENT APPLICATION NUMBER: US/10/659,800  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: 10/040,315  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/107,771  
; PRIOR FILING DATE: 1998-11-09  
; PRIOR APPLICATION NUMBER: PCT/US98/17883  
; PRIOR FILING DATE: 1998-08-28  
; PRIOR APPLICATION NUMBER: 09/103,754  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 09/339,472  
; PRIOR FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 488  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-659-800-6

Query Match 100.0%; Score 2594; DB 4; Length 488;  
Best Local Similarity 100.0%; Pred. No. 8.9e-233;  
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVVDAAAGPDVGAAGDAPAPAPKDGDAVG 60  
Db 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVVDAAAGPDVGAAGDAPAPAPKDGDAVG 60  
QY 61 SGHWE LRCHRLQDSLFSDDSGFSNYRGILNMCVWMLISNARLFLENLIKGI LVDPIQV 120  
Db 61 SGHWE LRCHRLQDSLFSDDSGFSNYRGILNMCVWMLISNARLFLENLIKGI LVDPIQV 120  
QY 121 VSLFLKDPHWPAPCLVIAAVFAVAFAVEKRLAVGALTEQAGLLHVNALATILCPPA 180  
Db 121 VSLFLKDPHWPAPCLVIAAVFAVAFAVEKRLAVGALTEQAGLLHVNALATILCPPA 180  
QY 181 AVVLIVESTITVGSILMAHTILFLKLFYRDVNSWCRARAKAASAGKASSAAABHT 240  
Db 181 AVVLIVESTITVGSILMAHTILFLKLFYRDVNSWCRARAKAASAGKASSAAABHT 240  
QY 241 VSYPDNLTYYRDLYYFLFAPTLCELNPFPRSPRIKRFLLRRLLEMLFTQLOVGLIQWM 300  
Db 241 VSYPDNLTYYRDLYYFLFAPTLCELNPFPRSPRIKRFLLRRLLEMLFTQLOVGLIQWM 300  
QY 301 VPTIIONSKPKMDYSRIIRLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOPGDBEF 360  
Db 301 VPTIIONSKPKMDYSRIIRLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOPGDBEF 360  
QY 361 YRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVS 420  
Db 361 YRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVS 420  
QY 421 VPLMFRLMAFTGMAAOIPLAMFVGRFQGNYGNAAVWLSLIIQPIAVLWYVHDYYVLN 480  
Db 421 VPLMFRLMAFTGMAAOIPLAMFVGRFQGNYGNAAVWLSLIIQPIAVLWYVHDYYVLN 480  
QY 481 YEAPAAEA 488

Db 481 YEAPAAEA 488

RESULT 5  
US-11-117-005-5  
; Sequence 5, Application US/11117005  
; Publication No. US20050193446A1  
; GENERAL INFORMATION:  
; APPLICANT: Zou, Utao  
; APPLICANT: Taylor, David C  
; APPLICANT: Wei, Yangdou  
; APPLICANT: Jako, Colette C  
; TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants  
; FILE REFERENCE: 3015-5684.1US  
; CURRENT APPLICATION NUMBER: US/11/117,005  
; CURRENT FILING DATE: 2005-04-28  
; PRIOR APPLICATION NUMBER: 09/623,514  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: PCT/CA99/01202  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: 60/112,812  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 488  
; TYPE: PRT  
; ORGANISM: human  
US-11-117-005-5

Query Match 100.0%; Score 2594; DB 6; Length 488;  
Best Local Similarity 100.0%; Pred. No. 8.9e-233;  
Matches 488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVVDAAAGPDVGAAGDAPAPAPKDGDAVG 60  
Db 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVVDAAAGPDVGAAGDAPAPAPKDGDAVG 60  
QY 61 SGHWE LRCHRLQDSLFSDDSGFSNYRGILNMCVWMLISNARLFLENLIKGI LVDPIQV 120  
Db 61 SGHWE LRCHRLQDSLFSDDSGFSNYRGILNMCVWMLISNARLFLENLIKGI LVDPIQV 120  
QY 121 VSLFLKDPHWPAPCLVIAAVFAVAFAVEKRLAVGALTEQAGLLHVNALATILCPPA 180  
Db 121 VSLFLKDPHWPAPCLVIAAVFAVAFAVEKRLAVGALTEQAGLLHVNALATILCPPA 180  
QY 181 AVVLIVESTITVGSILMAHTILFLKLFYRDVNSWCRARAKAASAGKASSAAABHT 240  
Db 181 AVVLIVESTITVGSILMAHTILFLKLFYRDVNSWCRARAKAASAGKASSAAABHT 240  
QY 241 VSYPDNLTYYRDLYYFLFAPTLCELNPFPRSPRIKRFLLRRLLEMLFTQLOVGLIQWM 300  
Db 241 VSYPDNLTYYRDLYYFLFAPTLCELNPFPRSPRIKRFLLRRLLEMLFTQLOVGLIQWM 300  
QY 301 VPTIIONSKPKMDYSRIIRLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOPGDBEF 360  
Db 301 VPTIIONSKPKMDYSRIIRLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOPGDBEF 360  
QY 361 YRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVS 420  
Db 361 YRDWNNSESVTYFMQNNNI PVHKWCIRHFYKPMRLRGSSKMMARTGVFLASAFHEYLVS 420  
QY 421 VPLMFRLMAFTGMAAOIPLAMFVGRFQGNYGNAAVWLSLIIQPIAVLWYVHDYYVLN 480  
Db 421 VPLMFRLMAFTGMAAOIPLAMFVGRFQGNYGNAAVWLSLIIQPIAVLWYVHDYYVLN 480  
QY 481 YEAPAAEA 488  
Db 481 YEAPAAEA 488

RESULT 6

```
US-10-741-600-1339
; Sequence 1339, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1339
; LENGTH: 488
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-741-600-1339

Query Match          99.8%; Score 2588; DB 5; Length 488;
Best Local Similarity 99.8%; Pred. No. 3.2e-232;
Matches 487; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60
DB 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60
QY 61 SGHWELRCHRLQDSLFFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP1QV 120
DB 61 SGHWELRCHRLQDSLFFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP1QV 120
QY 121 VSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTQAGLLHVANLATILCFPA 180
DB 121 VSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTQAGLLHVANLATILCFPA 180
QY 181 AVVLVSEITPVGSLALMAHTILFLKLFSTRDVSNCRRARAKASAGKASAAAPT 240
DB 181 AVVLVSEITPVGSLALMAHTILFLKLFSTRDVSNCRRARAKASAGKASAAAPT 240
QY 241 VSYPDNLTYRDLVYFLFAPFLCYELNPPRSRIRKRLRILEMFFTOLOVGLIQQM 300
DB 241 VSYPDNLTYRDLVYFLFAPFLCYELNPPRSRIRKRLRILEMFFTOLOVGLIQQM 300
QY 301 VPTIÖNSMKPKMDYSRIIERLLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOFGDREF 360
DB 301 VPTIÖNSMKPKMDYSRIIERLLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOFGDREF 360
QY 361 YRDWNNSESYTYFWQNNNIIVHKWCIHFYKPMLRSSKMMARTGVFLASAFHEHYLS 420
DB 361 YRDWNNSESYTYFWQNNNIIVHKWCIHFYKPMLRSSKMMARTGVFLASAFHEHYLS 420
QY 421 VPLMFRLMAFTGMAQIPLAMPVGRFQGNYGNAAVMLSLIGQPIAVLMYVHDYVYLN 480
DB 421 VPLMFRLMAFTGMAQIPLAMPVGRFQGNYGNAAVMLSLIGQPIAVLMYVHDYVYLN 480
QY 481 YEAPAAEA 488
DB 481 YEAPAAEA 488

RESULT 7
US-10-741-600-1341
; Sequence 1341, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1341
```

```
; LENGTH: 488
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-741-600-1341

Query Match          99.8%; Score 2588; DB 5; Length 488;
Best Local Similarity 99.8%; Pred. No. 3.2e-232;
Matches 487; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60
DB 1 MGDGSSRRRTGSRPSSHGGGPPAAAEVYRDAAAGPDVGAAGDAPAPAPNKGDAVG 60
QY 61 SGHWELRCHRLQDSLFFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP1QV 120
DB 61 SGHWELRCHRLQDSLFFSSDSGFSNYRGILNMCVWMLILSNARLPLENLIKXGILVDP1QV 120
QY 121 VSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTQAGLLHVANLATILCFPA 180
DB 121 VSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTQAGLLHVANLATILCFPA 180
QY 181 AVVLVSEITPVGSLALMAHTILFLKLFSTRDVSNCRRARAKASAGKASAAAPT 240
DB 181 AVVLVSEITPVGSLALMAHTILFLKLFSTRDVSNCRRARAKASAGKASAAAPT 240
QY 241 VSYPDNLTYRDLVYFLFAPFLCYELNPPRSRIRKRLRILEMFFTOLOVGLIQQM 300
DB 241 VSYPDNLTYRDLVYFLFAPFLCYELNPPRSRIRKRLRILEMFFTOLOVGLIQQM 300
QY 301 VPTIÖNSMKPKMDYSRIIERLLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOFGDREF 360
DB 301 VPTIÖNSMKPKMDYSRIIERLLKLAVPNHLIWLIFPYWLPHSCLNAVAELMOFGDREF 360
QY 361 YRDWNNSESYTYFWQNNNIIVHKWCIHFYKPMLRSSKMMARTGVFLASAFHEHYLS 420
DB 361 YRDWNNSESYTYFWQNNNIIVHKWCIHFYKPMLRSSKMMARTGVFLASAFHEHYLS 420
QY 421 VPLMFRLMAFTGMAQIPLAMPVGRFQGNYGNAAVMLSLIGQPIAVLMYVHDYVYLN 480
DB 421 VPLMFRLMAFTGMAQIPLAMPVGRFQGNYGNAAVMLSLIGQPIAVLMYVHDYVYLN 480
QY 481 YEAPAAEA 488
DB 481 YEAPAAEA 488

RESULT 8
US-10-157-855-17
; Sequence 17, Application US/10157855
; Publication No. US20020170091A1
; GENERAL INFORMATION:
; APPLICANT: Lasener, Michael W.
; APPLICANT: Ruzsineky, Diane M.
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE REFERENCE: 16516.158
; CURRENT APPLICATION NUMBER: US/10/157,855
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 09/326,203
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 500
; TYPE: PR
; ORGANISM: Rattus sp.
US-10-157-855-17

Query Match          88.6%; Score 2299.5; DB 4; Length 500;
```

Beet Local Similarity 85.9%; Pred. No. 2.7e-205;  
Matches 427; Conservative 26; Mismatches 31; Indels 13; Gaps 3;

```
Qy 1 MGBR---GSSRRRRRGSSRSSHGGGPPAAAEEDVDAAGPDVGAAGPAPAP----- 51
Db 1 MGBRGAGSSRRRRRGSSRSVIGSGGPMVDEEEDVDAAGPDGAGGDPAPAPAPAH 60
Qy 52 --NKDGDAGVSGHWEIACHRLQDSLFSDSGFSNYRGILMNCVWMLISNARLFLENLI 109
Db 61 TRDKRQRTSVGDSGHWEIACHRLQDSLFSDSGFSNYRGILMNCVWMLISNARLFLENLI 120
Qy 110 KYGILVDPIDQVSLFLKDPHSPAPCLVIAANVFAVAQVEKRLAVGALTQAGILLHV 169
Db 121 KYGILVDPIDQVSLFLKDPYSPAPCLTIASNIPIVATFOIEKRLSVGALTQMGILLHV 180
Qy 170 ANLATILCFPAVALVLESITPVGSLFALASYSIIFLKLFSYRDVNLWCQRQRRVAKAVS 240
Db 181 VNLATILCFPAVALVLESITPVGSLFALASYSIIFLKLFSYRDVNLWCQRQRRVAKAVS 240
Qy 228 AGKASASAAAPHTVSYPNLTYYRDIYFLFAPTLCEYELNFPSPRIIRKFLRLILEMLF 287
Db 241 AGKAVSGAAQNTVSYPNLTYYRDIYFLFAPTLCEYELNFPSPRIIRKFLRLILEMLF 300
Qy 288 FTQLOVGLIQOMMVPTIONSMPKPKMDYSRIIEBLKLAVPNHLIMLIFPYWLFHSCIN 347
Db 301 FTQLOVGLIQOMMVPTIONSMPKPKMDYSRIIEBLKLAVPNHLIMLIFPYWLFHSCIN 360
Qy 348 AVAELMOFGDREFFYDMNNSVTFWQNMNIPVHKCIRHFYKEMLRGSSKMMARIGV 407
Db 361 AVAELLQFGDREFFYDMNNSVTFWQNMNIPVHKCIRHFYKEMLRGSSKMMARIGV 420
Qy 408 FLASAFHEIYVSVDLRFRLMAFTGMAQIPLAMFVGRFQNGYGNAAVWLSLIIGQPI 467
Db 421 FLASAFHEIYVSIDLRFRLMAFTGMAQIPLAMFVGRFQNGYGNAAVWLSLIIGQPI 480
Qy 468 AVIMYVHDYVYLVNYPAP 484
Db 481 AVIMYVHDYVYLVNYPAP 497

RESULT 9
US-10-647-517-24
; Sequence 24, Application US/10647517
; Publication No. US20050102716A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING ALTERED LEVELS OF STEROL
; FILE REFERENCE: MTC 6462.1
; CURRENT APPLICATION NUMBER: US/10/647,517
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US/09/548,256
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/128,995
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 500
; TYPE: PRT
; ORGANISM: rat
US-10-647-517-24
```

Query Match 88.6%; Score 2299.5; DB 5; Length 500;  
Beet Local Similarity 85.9%; Pred. No. 2.7e-205;  
Matches 427; Conservative 26; Mismatches 31; Indels 13; Gaps 3;

```
Qy 1 MGBR---GSSRRRRRGSSRSSHGGGPPAAAEEDVDAAGPDVGAAGPAPAP----- 51
Db 1 MGBRGAGSSRRRRRGSSRSVIGSGGPMVDEEEDVDAAGPDGAGGDPAPAPAPAH 60
Qy 52 --NKDGDAGVSGHWEIACHRLQDSLFSDSGFSNYRGILMNCVWMLISNARLFLENLI 109
Db 1 MGBRGAGSSRRRRRGSSRSVIGSGGPMVDEEEDVDAAGPDGAGGDPAPAPAPAH 60
```

```
Db 61 TRDKRQRTSVGDSGHWEIACHRLQDSLFSDSGFSNYRGILMNCVWMLISNARLFLENLI 120
Qy 110 KYGILVDPIDQVSLFLKDPHSPAPCLVIAANVFAVAQVEKRLAVGALTQAGILLHV 169
Db 121 KYGILVDPIDQVSLFLKDPYSPAPCLTIASNIPIVATFOIEKRLSVGALTQMGILLHV 180
Qy 170 ANLATILCFPAVALVLESITPVGSLFALASYSIIFLKLFSYRDVNLWCQRQRRVAKAVS 240
Db 181 VNLATILCFPAVALVLESITPVGSLFALASYSIIFLKLFSYRDVNLWCQRQRRVAKAVS 240
Qy 228 AGKASASAAAPHTVSYPNLTYYRDIYFLFAPTLCEYELNFPSPRIIRKFLRLILEMLF 287
Db 241 AGKAVSGAAQNTVSYPNLTYYRDIYFLFAPTLCEYELNFPSPRIIRKFLRLILEMLF 300
Qy 288 FTQLOVGLIQOMMVPTIONSMPKPKMDYSRIIEBLKLAVPNHLIMLIFPYWLFHSCIN 347
Db 301 FTQLOVGLIQOMMVPTIONSMPKPKMDYSRIIEBLKLAVPNHLIMLIFPYWLFHSCIN 360
Qy 348 AVAELMOFGDREFFYDMNNSVTFWQNMNIPVHKCIRHFYKEMLRGSSKMMARIGV 407
Db 361 AVAELLQFGDREFFYDMNNSVTFWQNMNIPVHKCIRHFYKEMLRGSSKMMARIGV 420
Qy 408 FLASAFHEIYVSVDLRFRLMAFTGMAQIPLAMFVGRFQNGYGNAAVWLSLIIGQPI 467
Db 421 FLASAFHEIYVSIDLRFRLMAFTGMAQIPLAMFVGRFQNGYGNAAVWLSLIIGQPI 480
Qy 468 AVIMYVHDYVYLVNYPAP 484
Db 481 AVIMYVHDYVYLVNYPAP 497
```

```
RESULT 10
US-10-415-620-2
; Sequence 2, Application US/10415620
; Publication No. US20040076977A1
; GENERAL INFORMATION:
; APPLICANT: REID, SUZANNE J
; APPLICANT: FORD, CHRISTINE A
; APPLICANT: GEORGES, MICHEL A J
; APPLICANT: COPPIETERS, WOUTER H R
; APPLICANT: GRISART, BERNARD M J J
; APPLICANT: SNEEL, RUSSELL G
; APPLICANT: SPELMAN, RICHARD J
; TITLE OF INVENTION: MARKER ASSISTED SELECTION OF BOVINE FOR IMPROVED MILK COMPOSITION
; FILE REFERENCE: P452508 CIE
; CURRENT APPLICATION NUMBER: US/10/415,620
; CURRENT FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9434)..(9434)
; OTHER INFORMATION: ambiguous nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9436)..(9496)
; OTHER INFORMATION: ambiguous nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10402)..(10417)
; OTHER INFORMATION: ambiguous nucleotides
US-10-415-620-2
```

Query Match 88.4%; Score 2292.5; DB 4; Length 489;  
Beet Local Similarity 88.0%; Pred. No. 1.2e-204;

Matches 434; Conservative 15; Mismatches 31; Indels 13; Gaps 3;

```
OY 1 MGDR---GSSRRRTGSRPSSHGCGPAAAEVYRDAAGPVGAGDAPAPAPBNKGDGA 57
    |||
    1 MGDGAGAGSGRRRTGSRPSTIGGSGPAAAEVYR-----DVAGGDA PVRD TDKGDV 54
OY 58 GVGSGHWE LRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKXGILVDP 117
    |||
    55 DVGSGHMLRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKXGILVDP 114
OY 118 IQVNSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTBOAGLLHGVNLTATLTC 177
    |||
    115 IQVNSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTBOAGLLHGVNLTATLTC 174
OY 178 FPAAVVLLVESITTPVGSLLALMAHTIIFLKLFSYRDVNSWCCR-----ARAKASAGKXAS 233
    |||
    175 FPAAVVLLVESITTPVGSLLALMAHTIIFLKLFSYRDVNSWCCR-----ARAKASAGKXAS 234
OY 234 SAAPHVTSYDPDNLTYRDLYFLFAPFLCYELNPPRSRIRKRLRLRILEMLFFTOLOV 293
    |||
    235 GGAAGRTVSYDPDNLTYRDLYFLFAPFLCYELNPPRSRIRKRLRLRILEMLFFTOLOV 294
OY 294 GLIQOMWPTTQNSMKPEKMDYSRIERLLKLA VPHNLWLIFPYWLPHSCLNAVAELM 353
    |||
    295 GLIQOMWPTTQNSMKPEKMDYSRIERLLKLA VPHNLWLIFPYWLPHSCLNAVAELM 354
OY 354 QGDRREFYRDWNNSESYTFWQNNNIPVHKCIRHFYKPMRLRSGSSKMAARTGVFLASAF 413
    |||
    355 QGDRREFYRDWNNSESYTFWQNNNIPVHKCIRHFYKPMRLRSGSSKMAARTGVFLASAF 414
OY 414 FHEYLVSPLRMFLMAFTGMAOIPLAWFYGRFQNGYGAAYVLSLITGPIAVLMYV 473
    |||
    415 FHEYLVSPLRMFLMAFTGMAOIPLAWFYGRFQNGYGAAYVLSLITGPIAVLMYV 474
OY 474 HDYVLYNREAPAA 486
    |||
    475 HDYVLYNREAPAA 487
Db
```

## RESULT 11

```
US-10-415-620-5
; Sequence 5, Application US/10415620
; Publication No. US20040076977A1
; GENERAL INFORMATION:
; APPLICANT: REID, SUZANNE J
; APPLICANT: FORD, CHRISTINE A
; APPLICANT: GEORGES, MICHEL A J
; APPLICANT: COPIETERS, WOUTER H R
; APPLICANT: GRISART, BERNARD M J J
; APPLICANT: SNEEL, RUSSELL G
; APPLICANT: SPELTMAN, RICHARD J
; TITLE OF INVENTION: MARKER ASSISTED SELECTION OF BOVINE FOR IMPROVED MILK COMPOSITION
; FILE REFERENCE: P452508 CJE
; CURRENT APPLICATION NUMBER: US/10/415,620
; CURRENT FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Bos taurus
; NAME/KEY: VARIANT
; LOCATION: (232)..(232)
; OTHER INFORMATION: an amino acid substitution (K -> A) caused by a polymorphism at b
; OTHER INFORMATION: residues 7224-7225 of the genomic sequence (measured from the adenine
; OTHER INFORMATION: residue of the translation start codon). Lysine (K) corresponds
; OTHER INFORMATION: to the Q allele, alanine (A) corresponds to the q allele.
US-10-415-620-5
```

Query Match 88.4%; Score 2292.5; DB 4; Length 489;  
Best Local Similarity 88.0%; Pred. No. 1.2e-204;  
Matches 434; Conservative 15; Mismatches 31; Indels 13; Gaps 3;

```
OY 1 MGDR---GSSRRRTGSRPSSHGCGPAAAEVYRDAAGPVGAGDAPAPAPBNKGDGA 57
    |||
    1 MGDGAGAGSGRRRTGSRPSTIGGSGPAAAEVYR-----DVAGGDA PVRD TDKGDV 54
OY 58 GVGSGHWE LRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKXGILVDP 117
    |||
    55 DVGSGHMLRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKXGILVDP 114
OY 118 IQVNSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTBOAGLLHGVNLTATLTC 177
    |||
    115 IQVNSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTBOAGLLHGVNLTATLTC 174
OY 178 FPAAVVLLVESITTPVGSLLALMAHTIIFLKLFSYRDVNSWCCR-----ARAKASAGKXAS 233
    |||
    175 FPAAVVLLVESITTPVGSLLALMAHTIIFLKLFSYRDVNSWCCR-----ARAKASAGKXAS 234
OY 234 SAAPHVTSYDPDNLTYRDLYFLFAPFLCYELNPPRSRIRKRLRLRILEMLFFTOLOV 293
    |||
    235 GGAAGRTVSYDPDNLTYRDLYFLFAPFLCYELNPPRSRIRKRLRLRILEMLFFTOLOV 294
OY 294 GLIQOMWPTTQNSMKPEKMDYSRIERLLKLA VPHNLWLIFPYWLPHSCLNAVAELM 353
    |||
    295 GLIQOMWPTTQNSMKPEKMDYSRIERLLKLA VPHNLWLIFPYWLPHSCLNAVAELM 354
OY 354 QGDRREFYRDWNNSESYTFWQNNNIPVHKCIRHFYKPMRLRSGSSKMAARTGVFLASAF 413
    |||
    355 QGDRREFYRDWNNSESYTFWQNNNIPVHKCIRHFYKPMRLRSGSSKMAARTGVFLASAF 414
OY 414 FHEYLVSPLRMFLMAFTGMAOIPLAWFYGRFQNGYGAAYVLSLITGPIAVLMYV 473
    |||
    415 FHEYLVSPLRMFLMAFTGMAOIPLAWFYGRFQNGYGAAYVLSLITGPIAVLMYV 474
OY 474 HDYVLYNREAPAA 486
    |||
    475 HDYVLYNREAPAA 487
Db
```

## RESULT 12

```
US-10-482-936-4
; Sequence 4, Application US/10482936
; Publication No. US20040234986A1
; GENERAL INFORMATION:
; APPLICANT: Arbeitsgemeinschaft Deutscher Rindzuechter e.V.
; TITLE OF INVENTION: Method of testing a mammal for its predisposition for
; TITLE OF INVENTION: fat content of milk and/or its predisposition for meat
; FILE REFERENCE: F 1078 BP
; CURRENT APPLICATION NUMBER: US/10/482,936
; CURRENT FILING DATE: 2004-04-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-482-936-4
```

Query Match 87.8%; Score 2278.5; DB 5; Length 489;  
Best Local Similarity 87.8%; Pred. No. 2.3e-203;  
Matches 433; Conservative 14; Mismatches 33; Indels 13; Gaps 3;

```
OY 1 MGDR---GSSRRRTGSRPSSHGCGPAAAEVYRDAAGPVGAGDAPAPAPBNKGDGA 57
    |||
    1 MGDGAGAGSGRRRTGSRPSTIGGSGPAAAEVYR-----DVAGGDA PVRD TDKGDV 54
OY 58 GVGSGHWE LRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKXGILVDP 117
    |||
    55 DVGSGHMLRCHRLQDSLFSSDSGFSNRYGILNMCVWMLISNARLPLENLIKXGILVDP 114
OY 118 IQVNSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTBOAGLLHGVNLTATLTC 177
    |||
    115 IQVNSLFLKDPHSPAPCLVIAANVFAVAAFQVEKRLAVGALTBOAGLLHGVNLTATLTC 174
Db
```





```
Db 241 KKVSGAAAOQAVSYPDNLTYRDLYYFIAPFLCYELNFPSPRIKQFLRRVLEMLFFT 300
Qy 290 QLOVGLIQOMVPTIQNSMKPFKMDYSRIIERLLKLAVPNHLIWLIFPYLPHSCINAV 349
Db 301 QLOVGLIQOMVPTIHSNMPKFKMDYSRIIERLLKLAVPNHLIWLIFPYLPHSCINAV 360
Qy 350 AELMOFGDREFYRDMMNSSEVTYFWQNNIPVHKCIRHFYKPMRLRGSSKMMARTGVFL 409
Db 361 AELMOFGDREFYRDMMNSSEVTYFWQNNIPVHKCIRHFYKPMRLRGSSKMMARTGVFL 420
Qy 410 ASAFHEHYLVSVPLRMFLMAFTGMAQIPLAMFVGRFFQGNYGNAAVMLSLIIGPIAV 469
Db 421 TSAFFHEHYLVSVPLRMFLMAFTGMAQIPLAMFVGRFFQGNYGNAAVMLSLIIGPIAV 480
Qy 470 LMYVHDYVLYNLYEAP 484
Db 481 LMYVHDYVLYNLYDAP 495
```

```
RESULT 15
US-10-273-438-10
; Sequence 10, Application US/10273438
; Publication NO. US20030072157A1
; GENERAL INFORMATION:
; APPLICANT: Farese, Robert V.
; APPLICANT: Cages, Sylvaine
; APPLICANT: Smith, Steven
; APPLICANT: Erickson, Sandra
; TITLE OF INVENTION: Diacylglycerol O-Acyltransferase
; FILE REFERENCE: UCAL-105CIP2
; CURRENT APPLICATION NUMBER: US/10/273,438
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US/10/040,315
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/107,771
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: PCT/US98/17883
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/103,754
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 09/339,472
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 498
; TYPE: PRT
; ORGANISM: mus musculus
US-10-273-438-10
```

```
Query Match 87.6%; Score 2273.5; DB 4; Length 498;
Best Local Similarity 84.8%; Pred. No. 7e-203;
Matches 420; Conservative 29; Mismatches 35; Indels 11; Gaps 3;
```

```
Qy 1 MGDR---GSSRRRTGSRPSSHGSGGPAABEEVDAAAGPDVGAADAPAP----- 51
Db 1 MGDRGGAGSSRRRRIGSRVSVGGSGPRVEEDVDAVSPDLGAGDAPAPAPAHTR 60
Qy 52 NKGDAGVSGHMLRCHRLQDSLFSDSGFSNRYGILNWCVMILISNARILENLIKY 111
Db 61 DKDRSTVGDDGMDLRCHRLQDSLFSDSGFSNRYGILNWCVMILISNARILENLIKY 120
Qy 112 GILVDPIDVLSFLKDPHSWAPCLVIAANVAVAAPOVEKRLAVALTEQAGILLHYAN 171
Db 121 GILVDPIDVLSFLKDPHSWAPCLVIAANVAVAAPOVEKRLAVALTEQAGILLHYAN 180
Qy 172 LATICPPAAVVLVESTITPVGSLALMAHTLFLKLFSDVNSMC--RRARAKAASAG 229
Db 181 LATICPPAAVVALVESTITPVGSVALASYSIMFLKLYSRDVNLMCRGRKAKAVSTG 240
Qy 230 KKASSAAPHTVSYPDNLTYRDLYYFLPAFLCYELNFPSPRIKQFLRRVLEMLFFT 289
```

```
Db 241 KKVSGAAAOQAVSYPDNLTYRDLYYFIAPFLCYELNFPSPRIKQFLRRVLEMLFFT 300
Qy 290 QLOVGLIQOMVPTIQNSMKPFKMDYSRIIERLLKLAVPNHLIWLIFPYLPHSCINAV 349
Db 301 QLOVGLIQOMVPTIHSNMPKFKMDYSRIIERLLKLAVPNHLIWLIFPYLPHSCINAV 360
Qy 350 AELMOFGDREFYRDMMNSSEVTYFWQNNIPVHKCIRHFYKPMRLRGSSKMMARTGVFL 409
Db 361 AELMOFGDREFYRDMMNSSEVTYFWQNNIPVHKCIRHFYKPMRLRGSSKMMARTGVFL 420
Qy 410 ASAFHEHYLVSVPLRMFLMAFTGMAQIPLAMFVGRFFQGNYGNAAVMLSLIIGPIAV 469
Db 421 TSAFFHEHYLVSVPLRMFLMAFTGMAQIPLAMFVGRFFQGNYGNAAVMLSLIIGPIAV 480
Qy 470 LMYVHDYVLYNLYEAP 484
Db 481 LMYVHDYVLYNLYDAP 495
```

```
Search completed: May 5, 2006, 12:46:31
Job time : 57 secs
```